DOCUMENT RESUME

ED 161 885 TM 007 175

AUTHOR	Byrn, Darcie; And Others
TITLE	Evaluation in Extension.
SPONS AGENCY	Federal Extension Service (DOA), Washington, D.C.
PUB DATE	[67]
NOTE	119p.
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	Data Analysis; Data Collection; Data Processing; *Educational Assessment; Evaluation Criteria; *Evaluation Methods; *Extension Education; Formative Evaluation; *Guides; Interviews; Measurement Goals; *Program Evaluation; Research Methodology; *Rural Extension; Sampling; Technical Reports; Test Construction

ABSTRACT

The authors have written this manual to aid workers in the Cooperative Extension Service of the United States to be better able to understand and apply the principles and methods of evaluation. The manual contains three sections which cover the nature and place of evaluation in extension work, the evaluation process, and the uses of evaluation results. Within these sections, articles involve casual, self-checking, and do-it-yourself evaluation; identifying problems and ovjectives; the development and follow-through of a plan of work; stages and purpose of an evaluation plan; characteristics of scientific evaluation methods; collection and selection of data; evaluation criteria; analysis of a teaching plan; the depth of an evaluative study; sources of data; sampling procedures; interviewing competencies; constructing and using measurement instruments; tabulating, analyzing and interpreting data; preparing an evaluation report; and understanding and applying evaluation results. A sample outline used to evaluate an agricultural extension program, an example of an informal evaluation of in-service training meetings with extension workers, sources of assistance for the evaluator in extension education, and a list of references are appended. (Author/JAC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED161885

Evaluation in Extension

Authors:*

Darcie Byrn
Jewell G. Fessenden
Fred P. Frutchey
Gladys Gallup
Joseph L. Matthews
Ward F. Porter
J. Neil Raudabaugh
Laurel K. Sabrosky

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Chairman of Author Committee:

Fred P. Frutchey

Editor:

Darcie Byrn

* **Staff members of:**

Division of Extension Research and Training
Federal Extension Service
United States Department of Agriculture

TM007 175

MP-273

Evaluation in Extension

Prepared by Division of Extension Research and Training

FILE COPY
EXTENSION
DIVISION OF
RESEARCH & TRAINING

Authors:*

Darcie Byrn
Jewell G. Fessenden
Fred P. Frutchey
Gladys Gallup
Joseph L. Matthews
Ward F. Porter
J. Neil Raudabaugh
Laurel K. Sabrosky

Chairman of Author Committee:

Fred P. Frutchey

Editor:

Darcie Byrn

*** Staff members of:**

Division of Extension Research and Training
Federal Extension Service
United States Department of Agriculture

IVES
H. M. IVES & SONS, INC.
TOPEKA, KANSAS

FOREWORD

Extension workers need facts upon which to base their decisions in their Cooperative Extension work. The purpose of evaluation in Extension is to help obtain and interpret facts. The authors have written this manual to aid workers in the Cooperative Extension Service of the United States to be better able to understand and apply the principles and methods of evaluation.

Examples are drawn from and the applications are made to the Cooperative Extension Service, but principles and methods discussed have application to extension-type education in other parts of the world and to other kinds of informal education.

Extension evaluation means the use of the scientific approach in providing facts as a basis for making decisions, drawing conclusions, or forming judgments about the organization and conduct of extension work. Evaluation in Extension is to the extension worker what agricultural and home economics research is to farmers, distributors and homemakers.

All extension workers do some evaluation in their day-to-day work. They make decisions, draw conclusions and form judgments. The aim of this manual is to help them improve their evaluations and thereby improve their effectiveness in program planning, teaching methods, use of local leadership and organization for extension work.

"Evaluation" is a state of mind in which an extension worker wants better information about his work and uses the information to make his work more effective. It is an investment which produces progress and growth in achieving the objectives for which an organization has been set up.

The earlier editions of the manual have been used for inservice training purposes at Cooperative Extension workshops, conferences and summer schools. The preliminary edition (1956) has been used in other countries for similar purposes and has been translated into Spanish by the Instituto Inter-Americano de Ciencias Agricolas de la OEA, Turrialba, Costa Rica. The manual has served as a reference guide to Cooperative Extension workers engaged in evaluation studies and to State leaders of evaluation in Cooperative Extension.

In the present edition the authors have endeavored to present evaluation that will be most useful to teachers and students of evaluation.

GLADYS GALLUP, *Director*
DIVISION OF EXTENSION RESEARCH AND TRAINING
FEDERAL EXTENSION SERVICE

ACKNOWLEDGMENTS

The manual was first prepared by Mrs. Laurel K. Sabrosky, Dr. Gladys Gallup and Dr. J. Paul Leagans of Cornell University for use in the evaluation course at the Western Regional Extension Summer School in 1946 at Colorado State University. Later it was used as training material for other evaluation courses and workshops and as reference material on evaluation.

A revision was made in 1949 with additional members of the Division contributing. During 1954-56 a third edition, now out-of-print, was prepared and published by the Government Printing Office.

Since that time members of the Division who taught courses in evaluation at extension summer schools and the USDA Graduate School have revised the third edition chapters and developed new materials. These have resulted in the present edition of the manual.

Mr. Meredith C. Wilson contributed chapters to the third edition and as former Director of the Division gave it administrative support. Dr. Mary L. Collings helped to develop the philosophy and methods of evaluation reflected in the manual. Mr. C. Herman Welch, Jr., Dr. Woodsoon Fishback and Mrs. Amy Cowing contributed their suggestions in planning the manual and reviewing the manuscript.

Many Federal, State and county extension workers have made suggestions for improving the preliminary manual and in reviewing manuscripts of the present chapters.

The members of the Division wish to express deep appreciation to Dr. Ralph W. Tyler for his guidance in evaluation which has been the basis for the development of the work in extension evaluation.

FRED P. FRUTCHEY
Chairman of Author Committee

CONTENTS

THE NATURE AND PLACE OF EVALUATION IN EXTENSION WORK

Page

- I. EVALUATION—WHAT IT IS 1
  - What Is Evaluation? 2
  - Degrees of Evaluation 2
    - Casual Everyday Evaluations 2
    - Self-Checking Evaluations 2
    - Do-It-Yourself Evaluations 3
    - Extension Studies 3
    - Scientific Research 3
  - Evaluation for Everyone 3
  - Evaluations Can Be Improved 4
  - Tolerable Error 4
  - Summary 5
  - Reference 5
- II. EVALUATION IN EXTENSION EDUCATION 6
  - What Is Cooperative Extension Work 6
  - Basic Philosophy of Extension Education 6
  - Teaching—Learning Process 6
  - Extension Education Evaluation 7
  - Extension Education and Social Action 7
  - The Extension Teaching or Program Development Cycle 7
    - Identification of the Problem 8
    - Determination of Objectives 8
    - Development of a Plan of Work 8
    - Follow Through on the Plan of Work 8
    - Determination of Progress 8
  - Summary 8
  - References 9
- III. THE PLACE OF EVALUATION IN EXTENSION 10
  - Why Evaluate? 10
  - What Can We Gain from Evaluation 10
  - Stages of Evaluation 11
  - Purposes of Evaluation 12
  - Who Should Do Evaluation 12
  - Summary 12
  - References 12
- IV. THE SCIENTIFIC APPROACH AND EXTENSION EVALUATION 13
  - Custom and Tradition 13
  - Authority 13
  - Personal Experience and Intuition 13
  - Scientific Method 14
  - Characteristics of the Scientific Method 14
    - Factual 14
    - Analytical 14
    - Impartial 14
    - Reliable 14
    - Objective 14
  - Science and Scientific Method 15
  - Five Questions 15
    - What Do You Want to Know? 15
    - Where, and How Will You Get the Information? 16
    - Who Will Collect the Information? 16
    - How Will the Information Be Analyzed? 16
    - What Does It Mean? 16
  - An Evaluation Outline 16
    - Need for the Evaluation 16

Page

    - Purposes of the Evaluation 16
    - Questions to Be Answered by the Evaluation 16
    - Sources of the Information 16
    - Collecting the Information 16
    - Selecting or Constructing a Record Form 16
    - Analysis and Tabulation of the Data for Use 17
    - Interpreting, Reporting and Applying the Findings 17
  - Summary 17
  - Reference 17

THE EVALUATION PROCESS

- V. GOALS AND OBJECTIVES 18
  - Need for Objectives 18
  - Kinds of Objectives 18
  - Levels of Objectives 18
    - Objectives of Society 19
    - Objectives of a Specific Organization or Group 19
    - Objectives of the Extension Agent 19
    - Objectives of the Clientele 19
  - Stating Educational Objectives 19
  - Criteria for Judging Extension Educational Objectives 20
  - Steps in Setting Goals and Objectives 21
  - Summary 21
  - References 21
- VI. THE PLACE OF TEACHING PLAN ANALYSIS IN THE EVALUATION PROCESS 22
  - Learning Experiences 22
  - Analysis of the Teaching Plan 23
  - Relationship Between Judging Learning Experiences and Educational Evaluation 24
  - Summary 24
  - References 24
- VII. EVIDENCE OF PROGRESS TOWARD OBJECTIVES 25
  - Evidence 25
  - Depth of Evaluation 25
    - In Terms of Changes in Behavior of People. 25
    - In Terms of Opportunity 25
  - Levels of Behavior Change at Which to Measure 26
  - Factors for Deciding Which Evidences of Behavior to Use 26
    - Which Evidences Will the Learners Have an Opportunity to Display? 26
    - Will the Things You Can Count Be Adequate as Evidence? 27
    - What Is Your Own Facility in Measuring Different Types of Behavior? 27
    - How Will You Limit the Number of Evidences to Look For? 27
    - What Will You Do About the Time Factor? 27
    - What Decisions Must You Make Regarding Units? 28
    - What Face Data Will You Need? 28
  - Summary 29
  - References 29

Page

VIII. IDENTIFYING PROBLEM AREAS AND STUDY OBJECTIVES 30
  The Importance of Problem Identification 30
  Selecting a Problem for Study 30
  Selecting and Defining Study Objectives 31
  Stating Study Objectives 31
  Summary 32
  References 32

IX. SOURCES OF DATA 33
  Primary or Secondary Source 33
  Useful Data Sources for the Extension Worker 34
    General 34
    Home Economics 35
    Extension 35
  Summary 36

X. SAMPLING 37
  General Definitions and Principles 37
    Sampling 37
    Population 37
    Representativeness 38
  Sample Size 38
  The Representative Sample 40
    List Sampling 40
    Area Sampling 41
    Grid or Random Point Sampling 42
    Group Sampling 42
    Stratified Sampling 42
  Judgment Sampling 43
  Limiting the Population to Study Size 43
  Testing the Sample 44
  Summary 44
  References 44

XI. METHODS OF COLLECTING DATA 45
  Methods of Collecting Data 45
  Considerations Which Influence Selection of Method 45
  Analysis of Each Method 45
    The Mail Questionnaire 45
    Personal Interview Schedules 46
    Distributed Questionnaires or Checklists 46
    Group Interviews 47
    Case Studies 47
    Systematic Observation Procedure 47
    Systematic Study of Available Records 48
  Planning to Carry Out the Evaluation 48
    Who Will Be Involved in Gathering the Information? 48
    When Will the Information Be Collected? 48
    Where Will the Study Be Made? 48
  Summary 48
  References 48

XII. INTERVIEWING 49
  Major Kinds of Interviews 49
    Face-to-Face Interview 49
    Telephone Interview 50
    Group Interview 50
  Responsibility of the Interviewer 50
    Maintain a Neat Personal Appearance 51
    Locate the Respondent, Get the Facts and Record Them 51
    Be Considerate and Honest with Respondents 51
    Understand the Purposes of the Study 51
    Be Thoroughly Familiar with the Instruments 51
    Follow Sampling Instructions 51
    Ask Questions Exactly as Written 51
    Check Work for Completeness (Editing) 51
  Selecting and Training Interviewers 52

Page

  Agenda for Training Interviewers 52
  Interviewing Procedures and Techniques 53
    Establishing Rapport 53
    Securing Information 54
    Closing the Interview 55
    Editing 55
    Interviewing Without a Schedule 55
  Summary 56

XIII. DEVICES FOR COLLECTING DATA 57
  The Need for Measuring Devices 57
  The Selection of the Measuring Device 57
  Some Illustrations of Devices 57
    Knowledge Questions or Checks on Levels of Information 57
    Understanding Questions 59
    Skill or Performance Ratings 59
    Attitude Testing 61
    Interest Checks 61
    Confidence Testing 62
    Practices Adopted or Action Taken 62
    Practices Adopted and Methods Responsible 63
  Summary 63
  References 63

XIV. CONSTRUCTION OF AN EVALUATION DEVICE 64
  Criteria for Constructing a Measuring Device 64
    Validity 64
      Behavior Validity 64
      Subject-Matter Validity 64
    Reliability 64
    Objectivity 64
    Practicability 65
    Simplicity 65
  Types of Questions 65
    Dichotomous Questions 65
    Multiple Choice Questions 65
    Open End Questions 66
  Wording of Questions 66
  Face or Descriptive Data Questions 66
  Core Questions 67
  Sequence of Questions 67
    Logical Order 67
    Psychological Order 67
  Mechanical Set Up 67
    Four Parts to a Questionnaire 67
    Physical Arrangement 68
    Adequate Space 68
      Good Appearance 68
    Some Special Rules for Constructing the Mail Questionnaire 68
  Pretesting the Questionnaire 68
  Summary 68
  References 69

XV. TABULATION 70
  Planning for Tabulation 70
    Resources and Skills Available 70
    Information Needed for Analyses 70
    Relationships to Be Studied 71
    Comparisons to Be Made 71
  Tabulation Techniques 71
    Tally Sheet 71
    Handsorting 72
    Tabulation Sheet 73
    Machine Tabulation 73
    Manual Tabulation Using Punched Cards 74
    Free-Answer Tabulation 74
  Summary 74
  References 75

Page

- XVI. ANALYSIS AND INTERPRETATION ... 76
  - Techniques of Analysis ... 76
    - Numerical Counts ... 76
    - Percentages ... 77
      - Correct Use of Base ... 77
      - Adding Percentages ... 77
      - Averaging Percentages ... 78
      - Rounding Percentages ... 78
    - Measures of Central Tendency ... 78
      - Arithmetic Mean ... 78
      - Mode ... 79
      - Median ... 79
    - Measures of Dispersion ... 79
    - Ranking ... 79
  - Interpretation Techniques ... 80
    - Generalization ... 80
    - Definition ... 80
    - Objectivity ... 80
    - Cause and Effect Relationships ... 81
    - Program Objectives ... 81
    - Potential ... 81
    - Trends ... 82
  - Summary ... 82
  - References ... 82
- XVII. PREPARING A RESEARCH OR STUDY REPORT ... 83
  - Planning a Study Report ... 83
    - Audience ... 83
    - Resources ... 84
  - Writing a Study Report ... 84
    - Report Outline ... 84
    - Organization and Content ... 84
      - Title ... 85
      - Summary Statement of Study ... 85
      - Table of Contents ... 85
      - Study Objectives and Procedures ... 85
      - Findings ... 85
      - Implications ... 85
      - Documentation ... 86
      - Appendix ... 86
    - Style of Writing ... 86
  - Methods of Presenting Data ... 86
    - Narrative Description ... 86
    - Tables ... 87
    - Graphs and Charts ... 87
      - Line Charts ... 87
      - Bar Charts ... 88
      - Pie Charts ... 88
      - Pictorial Charts ... 88
  - Summary ... 88
  - References ... 88

THE USES OF EVALUATION RESULTS

- XVIII. APPLYING THE RESULTS OF EVALUATION IN EXTENSION ... 89
  - Four Objectives ... 89
  - Three Ways to Improve Practices ... 89
    - Subjective Impressions ... 89
    - Professional Research ... 89
    - Evaluation ... 89
  - Basic Differences ... 89
    - Motivation ... 89
    - Working Conditions ... 89
  - Disadvantages of Calling in Professional Researchers ... 90
    - Impersonal ... 90
    - General ... 90
    - Detached ... 90
  - Getting Results ... 90
    - Involvement ... 90
  - Involving Others in Evaluation ... 90
    - When Deciding to Evaluate ... 91
    - While Planning the Evaluation ... 91
    - While Evaluation Is in Progress ... 91
    - When the Evaluation Is Completed ... 91
    - Afterward ... 92
  - Summary ... 92
  - References ... 92
- XIX. UNDERSTANDING AND USING EVALUATION REPORTS ... 93
  - Reports Must Communicate ... 93
  - Inhibitors to Communication ... 93
  - Options and Obligations ... 93
  - Personalized Presentation ... 94
  - Brevity ... 94
  - Three Operations ... 94
  - Appraising the Quality of a Report ... 94
    - Purpose ... 94
    - Method ... 94
    - Population ... 95
    - Sample ... 95
    - Timeliness ... 95
  - Summary ... 95
- XX. IN CONCLUSION ... 96
  - Why We Need Evaluation ... 96
  - Basic Elements of Program Evaluation ... 96
    - Pin-Pointing Objectives ... 96
    - Program Activities and Methods ... 96
    - Establishing Benchmarks ... 97
    - Determining Program Progress and Accomplishments ... 97
  - The Challenge ... 97

APPENDIX

- A. Evaluating a Pasture Improvement Program (Illustrative Outline) ... 98
- B. Example of an Informal Evaluation ... 102
- C. Professional Help for the Evaluator ... 103
- D. Evaluation References in Extension Service Review ... 106

THE NATURE AND PLACE OF EVALUATION IN EXTENSION WORK

Chapter I

EVALUATION – WHAT IT IS

Fred P. Frutchey

As County Agent Brown was driving home from a night meeting, he reflected on the meeting with a great deal of satisfaction. It was a fine meeting. He wished all his meetings were as good as this one.

He recalled the high level of interest which persisted throughout. He smiled contentedly as he remembered how he had to "kill the meeting rather than let it die." After the meeting many stayed on to ask questions and talk among themselves in small groups.

What a pleasant contrast to some other meetings! Meetings like this are what make a county agent's work mean something. You don't mind the lateness of the hour when you know your job was well done and that people were helped.

Under the glow of his satisfaction, however, Mr. Brown began to ask himself some questions. "They seemed enthusiastic enough, but what did they get out of the meeting? Was this an educational experience for them or was it purely entertainment? Did the meeting bring about any desired change in them? Or will it? What did I really have in mind to accomplish by the meeting? What were my objectives?"

Questions like these were analytical. Soil testing had been the topic. He hoped they would get some new information about soil testing and liming; that the meeting would create a favorable attitude toward soil testing and liming, and that they would test their soils and would lime when necessary.

Obviously, there could be no immediate evidence that they did test their soil and did apply recommended amounts of lime. He would check on that from time to time during the summer. However, there were several indications that they intended to act on the information he had given them.

He remembered how they "perked up their ears" when he showed them how easy it was to have the soil tested, and where they could get it done. They were equally attentive while he spoke about the value of liming, the effect of liming on yields, and the relation of cost of liming to increased value of crops raised. Some were busy making notes. Others raised questions which indicated they were applying the explanation to their own situations. Several asked him questions after the meeting.

It seemed to Mr. Brown that these observations indicated that many persons in the group had learned some of the things he was trying to explain to them. They had changed in that respect. From these observations he could conclude that most of them had a favorable attitude, and a fairly strong favorable attitude — so strong that many would probably do something about soil testing and liming. But that, of course, remained to be seen.

As Mr. Brown was thinking along these lines, he was evaluating. He never had any course or training in evaluation. The way he was thinking came natural to him.

As he drove along his thoughts moved on to why this meeting was so good and why some other meetings were not so good. What had he done this time to help bring about the changes in the people? More specifically, what had he done this time that was different from what he had done on those occasions which he would like to forget?

For one thing, there was that chart he and his secretary had worked on most of the afternoon. He remembered how the expressions of the audience changed as he went through his explanation, aided by the chart showing yields with and without adequate liming, relative costs and net gain.

He recalled, also, that this was the first time he had bothered to make a demonstration. When he had finished showing how easy it is to take a soil sample, he saw many of them taking down the address of the place to send the sample. He could tell he had their complete attention while he was relating how the results of the tests could be applied and what the effect on yield would be.

There were some other things which probably helped. For instance, he did much more careful planning and rehearsing for this meeting than he had done for other meetings. He felt more sure of himself at the meeting. Mr. Brown realized that all these things contributed to setting up a good learning situation for his audience. What he and others had done would influence the degree to which the people would change.

"So that is what my district agent meant when he said that a county agent's job is to create learning experiences by which people can change in the direction of educational objectives," thought Mr. Brown as he reached home and put away his car.

This story illustrates that, whether or not a person is trained in evaluation, he makes countless evaluations, anyway. Whether or not Mr. Brown was aware of it, his reflections on the trip home constitute a form of evaluation.

Evaluation is a common activity in which we all engage everyday. We finish a meal and evaluate it by saying, or only thinking, how good or bad it was. We may go further than evaluate the result; we may consider what was especially good or especially poor about it and why. The meat was tender and juicy, but the potatoes could have been cooked longer.

We read a book and react in some manner. It may be a very interesting one and we recommend it to our friends. It may leave us so "cold" that we do not bother to finish it.

What Is Evaluation?

The purpose of these illustrations is to point out that all of us engage in evaluation in our day-to-day affairs. Trifling indeed, is the action to which we do not apply the three steps in the evaluation process in some degree of intensity. First we make some observations or collect some information. Then we apply some standards or criteria to our observations. Finally, we form some judgment, draw some conclusion or make some decision.

These three elements are involved in all evaluation. It can be done so casually that we are hardly aware of doing it, such as looking out the window to decide whether or not to carry an umbrella. At the other extreme is scientific research in complicated problems to get information which people can use. Somewhere in between will fall most kinds of evaluation undertaken by extension personnel.

The better we do the evaluation job, the less likely we are to "jump" to conclusions which are misleading. The purpose of this manual is to discuss how the evaluation process can be used to aid us in drawing better conclusions and in making better decisions in our work. The manual is not designed to make the reader a specialist in extension research. Rather, it is intended to familiarize him with the processes of evaluation and to help him to make better evaluations in doing extension work.

Degrees of Evaluation

When we think of evaluation as a process of collecting information as a basis for making decisions, forming judgments and drawing conclusions, we realize it has much in common with scientific research. As with evaluation, we do scientific research to obtain usable information. Even so, there is a great difference between our casual everyday evaluations (the umbrella decision) and scientific research. The difference, however, is a matter of degree rather than kind. It lies in the difference in the degree to which scientific method is necessary in the solution of problems.

Degrees of anything are usually represented on a scale or continuum. Distance can be from zero inches to millions of miles. There is no gap in the scale. Temperature is also a continuous series represented in degrees from hot to cold. Other examples are age, young to old; weight, light to heavy; quality, as in adjectives of comparison, good-better-best; strength of attitudes, strong to weak and favorable to unfavorable; skill, degrees of skill; knowledge, amounts from none to much; sound, loudness of sound; and even color, wave lengths of light are expressed as degrees.

The examples illustrate that many characteristics of people are not on an "either-or" basis; not necessarily either present or absent; not either black or white. There are shades. A characteristic may be partly present and partly absent. The degrees of a characteristic blend into each other on a scale or continuum with no sharp lines of distinction.

In order to avoid emphasizing differences between evaluation and research, it is more useful to emphasize similarities. Casual everyday evaluations can be placed at one end of the scale and scientific research at the other end. All degrees are found between the two extremes.

Casual everyday evaluations	Scientific research

The value of such a scale is in realizing that we can improve our evaluations without becoming a scientist. Other locations can also be described on the scale between the two extremes, as follows:

Casual everyday evaluations	Self-checking evaluations	Do-it-yourself evaluations	Extension studies	Scientific research

The above locations are areas of the scale with no sharp lines of division. One blends into the other. The first three locations may be described as informal evaluations, which extension agents can do themselves.

Casual everyday evaluations need no further explanation. They are like the first impressions of County Agent Brown about his meeting or the umbrella decision. They are the ones we ordinarily make without much consideration of the principles of evaluation in the decisions we make about simple problems.

Self-checking evaluations represent a conscious attempt to apply principles of evaluation. Mr. Brown did some analytical thinking as he drove home from the meeting. He was doing some evaluation at a higher level. He was raising questions to himself about his initial feeling of satisfaction.

He considered to some extent the information he obtained by his observations and what conclusions he could draw from those observations.

Self-checking includes further checking on our ordinary observations. It includes talking further with others, writing to others for their judgment, sending out a brief questionnaire, having one filled out at a meeting and so on. Mr. Brown could do a little checking as he went around the country to see if any of the farmers had their soil tested since the meeting.

Do-it-yourself evaluations involve more planning and application of the principles of evaluation. They are more systematically done, more carefully planned and usually require some technical help. Each step in the evaluation is considered, planned and carried out with due consideration to evaluation principles.

These evaluations are not complex and involved. They are usually surveys which produce usable results and which can be done easily with some training in evaluation or with some technical help. The best of these sometimes pass as data for a master's thesis. They blend into the next location on the scale, "Extension studies."

The objective of a three-week summer school course in evaluation is that the participants reach the location on the scale of a do-it-yourself evaluation; that they plan and describe the steps in an evaluation which they can do themselves upon returning home.

The last two locations on the scale, "Extension studies" and "Scientific research" may be described as formal studies.

Extension studies are more involved and complicated to plan and carry out than any of the preceding locations on the scale. They are broader in scope. They require greater attention to sound principles of scientific procedure in order to secure the accuracy needed. Theses for master's degrees usually fall in this location.

Scientific research, of course is at the "top" of the scale, involving very complex problems and techniques for getting information from which conclusions can be drawn. Long time and experimental studies to determine cause and effect relationships are characteristic of this location. The invention of complicated techniques and procedures for getting information (for doing research) falls in this location. The top of the location may be illustrated by atomic research, satellite research or cancer research.

The scale of degrees of evaluation serves as a communications device in a discussion for identifying the kind of inquiry being considered and to be planned. It shows whether the inquiry is a very informal study or a very formal study. The kind of "research" under consideration can be identified on the continuum and misunderstanding cleared up. The design and techniques can then be adapted to what is valid and practical in the situation to obtain the accuracy needed.

Evaluation for Everyone

Extension workers have been hearing about evaluation during most of their association with Extension. They sometimes feel that as "doers" of extension work, evaluation and research are out of their field; that evaluation and research are something mystical or for persons with years of specialized training; that as extension teachers they have never done and will never do any evaluation or research.

The continuum of degrees of evaluation shows that there is a place for evaluation in every extension activity. Practically anyone can do evaluation at the first two levels with no help at all. Most of them can handle a do-it-yourself evaluation with a little counseling and they can cooperate at the other two levels.

The continuum also shows that we all commonly use methods of inquiry to get information as a basis for forming a judgment, making a decision, or drawing a conclusion. We can move up the scale and do better evaluations, but need not become full-time researchers at the far end of the scale.

Early in the days of the cooperation of the land-grant colleges and the U. S. Department of Agriculture in disseminating information relating to agriculture and home economics, it became evident that if extension work were to be effective facts would be needed about its clientele, personnel, program planning, teaching methods, local leadership and other factors.

The Cooperative Extension Service began to carry on studies which would provide information for making decisions about the conduct of extension teaching. These studies were more scientific than casual everyday evaluations and broader in scope.

It was recognized that since the Extension Service was recommending research findings to improve agriculture and homemaking, it must "practice what it preaches" and study itself through scientific procedures in improving its work. It was also recognized that if Extension carries on the study itself, the findings will be better understood and more likely to be used.

Any institution should plow back into its work a percentage of its funds to improve itself and make its work more effective. This is the function of evaluation in Extension.

As the size and importance of extension education has grown, more objective information was needed. More formal studies have been made and more personnel have been assigned to help extension agents to make better casual everyday evaluations and to make "do-it-yourself" evaluations of their teaching.

Evaluations Can Be Improved

Casual everyday evaluations can be misleading, and often are, because of errors in observation and interpretation. There can be considerable difference between what we see and what we think (interpret) we see. Great care must be taken to distinguish between what we actually see and what we "read into" what we see.

When we listen, for exanple, to a person speaking—and espeoially when we read a printed page—much of what we seem to hear or see is really supplied by our memory or what we expect to see or hear. We overlook misprints, assuming that we see the right letters, even though the wrong letters are actually before us. In this paragraph two misprints have been made intentionally. Did you notice them?

Our minds and our biases can influence our interpretations and cause errors, thereby affecting our judgments, decisions and conclusions.

Time and again it has been shown that persons having fleeting glimpses of behavior or during the excitment of the moment have made sworn statements of what they saw or heard only to have been proven positively wrong by further evidence.

Casual everyday evaluation can be improved by trying to avoid these errors. Some of the ways we can fortify ourselves and improve our evaluations are as follows:

1. Develop an attitude of mind by which we deliberately try to prove our interpretation wrong. Be critical of our own interpretations.
2. Recognize our individual subjective bias and take account of it in our interpretation.
3. Before drawing conclusions from a single instance, make further observations to determine how widespread or consistent the behavior is.
4. Check further to determine whether the behavior was due simply to courtesy, to make an impression, or was forced upon the person. Make allowances for or discount these things in our evaluation.
5. Consider other causes that could have been responsible for what we observed and check on these causes. More than one thing can cause an effect.
6. Be careful not to read into the observation what we hoped to find (wishful thinking) and overlook other reasonable interpretations.
7. In evaluating a result or product of behavior use criteria by which to judge the result or product, such as a score card.
8. Check further to determine whether a product made or things claimed to have been done by the person were really done by the person and not by someone else.
9. Recognize that many kinds of behavior are on a continuum, shading from one end of the scale to another, and are not on an all-or-none basis.
10. Determine whether the evidence we observed is actually the kind of evidence we would accept, or whether it cannot really be considered as evidence.
11. Finally, avoid jumping to conclusions, but if we must act (and that we must do many times) before all the evidence is in, act upon the best evidence available depending upon the seriousness of the consequence of our actions.

Tolerable Error

By following the preceding suggestions we can improve our evaluations. It would be unfortunate, however, if we go so far as to become perfectionists. Then we would not do any evaluation because all evaluations possess some error. Human beings are not infallible and cannot follow the preceding practices with perfection. Furthermore, absolute perfection is not necessary, and not practical.

Of course, in collecting information upon which to base a decision, make a judgment or draw a conclusion, it is necessary to collect "accurate" information and to make an "accurate" interpretation. But how accurate is "accurate"? No evaluation is perfectly accurate.

Casual everyday evaluations, however, are often subject to a high degree of inaccuracy. By improving our evaluations we reduce the size of the errors, at least to the point where we can tolerate the error that remains so we can make **more** accurate decisions.

In measuring the length of a room, a tape measure will give sufficient accuracy for purposes of laying carpet, or perhaps using a yardstick will be accurate enough. For some purposes merely "stepping off" the room is accurate enough.

A show of hands at meetings of dairy farmers may be accurate enough for some purposes to find out if the farmers are improving their pastures this year. For other purposes you may need a careful sample survey which would show **how** they are improving their pastures.

Suffice to say at this point that evaluation always involves some error but that better evaluations will reduce the error to a point where it is insignificant for practical use and will not mislead us nor others in making decisions, forming judgments and drawing conclusions.

Summary

The general concept of evaluation as a process of (1) collecting information, and (2) applying standards or criteria in (3) drawing conclusions, forming judgments, or making decisions, is useful because it can be applied to all kinds of situations confronting extension workers. All kinds of programs, methods, activities and situations in Extension can be evaluated.

We do not need to be research workers in order to use principles of evaluation. These principles can be applied by all persons. Perfect accuracy is not necessary; nor is it attainable, even in scientific research.

There is a wide range in degrees of evaluations from casual everyday evaluations to scientific research as indicated by the continuum.

We often make gross errors in our everyday evaluations and "jump" to conclusions. Care in the use of principles of evaluation can improve our judgments and decisions.

How far to go in applying evaluation principles depends upon the amount of error we can tolerate and still have useful results for our purposes.

As will be discussed in the next chapter, evaluation is not an isolated process. It is a continuing process used at all points on the extension cycle.

References

1. Good, C. V. and Scates, D. E. **Methods of Research.** New York, Appleton-Century-Crofts, 1954. Pages 1-12.
2. Appendix D—Evaluation References in **Extension Service Review.**

Chapter II

EVALUATION IN EXTENSION EDUCATION

J. Neil Raudabaugh

What Is Cooperative Extension Work?

The Cooperative Extension Service is a democratic educational arrangement among the people in the counties, the State land-grant colleges and universities, and the U. S. Department of Agriculture. It is legally referred to as Cooperative Extension Work in the basic legislation, the Smith-Lever Act, which was passed by Congress in 1914.

The purpose of Cooperative Extension Work, as stated in the original Act,[3] is " . . . to aid in diffusing among the people of the United States useful and practical information on subjects relating to agriculture and home economics and to encourage the application of the same." This Act further stipulates that " . . . agricultural extension work shall consist of the giving of instruction and practical demonstrations in agriculture and home economics to persons not attending or resident in said colleges in the several communities, and imparting to such persons information on said subjects through field demonstration, publication and otherwise." It is within the framework of this legislation, and other acts related to education, that the objectives and scope of cooperative extension educational programs are determined.

The basic administrative unit of the Cooperative Extension Service is the county. It is here that educational programs are planned and where extension teaching is done. The basic objective of the Cooperative Extension Service is to develop the people into understanding, effective, self-confident members of society who will have broad vision and be capable of meeting responsibilities in their homes, their communities, their State, and their nation. We in Extension express this objective in terms of developing people to the point where they, through their own initiative and ability, may effectively identify and solve the various problems directly affecting their welfare.

Some of the major fields of educational responsibility of the Cooperative Extension Service, in addition to efficient agricultural production and adjustment and home economics as such, are: Leadership development; community improvement; farm and home management; public affairs; marketing and distribution; conservation of natural resources; housing; health; and social relationships, adjustments and cultural values.[4]

Basic Philosophy of Extension Education

In cooperative extension education, it is acknowledged that people must be reached where they are; that is, at their present stage of educational development and experience and at their present level of interest and understanding. It is the function of Extension to teach people how to assess their own needs and how to solve their own problems, to help them acquire knowledge and understanding, and to inspire them to action. The extension teacher serves people by teaching them how to help themselves; in so doing, he teaches **people,** not subject matter. He teaches people **how** to think, not what to think.

The success of extension teaching depends on much more than knowledge and understanding of technical agriculture and/or home economics. It is equally important that the teachers understand people and their problems and that they know how to apply sound principles of education, psychology and sociology.

The goal of the extension teacher is to teach not only desirable practices and understandings, but also the interdependence of practices and understandings so fundamental to achieving satisfying family and community life. This is accomplished through programs which the people and extension personnel work out together.

It is the extension teacher's role, first, to inspire people as individuals, families and communities to work together to identify their own problems; second, to help them determine their objectives; and third, to counsel and supply technical help to them as needed in working toward these objectives.

Teaching-Learning Process

A brief consideration of education, and the elements in the processes of learning and teaching basic to the conduct of extension work, is helpful in understanding Extension as an educational process, as well as the place of evaluation in extension teaching. A definition of education which has been accepted generally by extension workers and which has been quoted often in extension publications is: "Education is the production of changes in human behavior."[11] It is the aim of extension education to influence people to make desirable changes in their behavior that will contribute to better farming and homemaking, to better family and community living.

The educational changes in behavior which concern extension teachers, and which are expressed in teaching level objectives, may be classified as follows:

1. Interests
2. Ideals and purposes
3. Information, knowledge and understanding
4. Abilities, skills, habits and practices
5. Attitudes and emotional responses

The steps in the learning process which contribute to the production of these behavioral changes have been identified and studied carefully. Beal and Bohlen, in their treatment of the diffusion process, suggest four steps which can be summarized in this manner:[2]

1. **Awareness**—People find out that the idea or practice exists.
2. **Interest**—People become concerned and recognize that the idea or practice has merit.
3. **Evaluation and Decision**—People make mental applications of the idea or practice to their own situations and reject it or decide to try it.
4. **Testing and Adoption**—People try the idea or practice and adopt it if they are satisfied.

Extension Education Evaluation

Effective educational programs provide for these steps in the learning or diffusion process and contribute to production of the previously mentioned kinds of behavioral changes. To evaluate a program, evidence must be collected about these kinds of behavior and behavioral changes in the people concerned. Educational evaluation, then, is the process of determining the degree to which desired behavioral changes have taken place or are currently taking place as a result of educational effort. Similarly, extension education evaluation is the process of determining how well desired behavioral changes have taken place or are taking place as a result of extension educational effort.

Extension Education and Social Action

Through systematic analysis of the action process of educational programs as they move forward, distinct and related action stages are suggested. These stages are involved in successful educational programs regardless of whether the educational approach used is intensive (individual or family centered) or extensive (group oriented). Extension educators are concerned about, (1) the areas in which there is need for change, (2) what direction the change should take, (3) how rapidly the change takes place, and (4) how change can be directed so that desired accomplishments are maximized.

There are many functions that must be performed in the successful and efficient conception and implementation of action oriented educational programs. These functions can be organized logically into a sequence of actions from the inception of an idea, or the recognition of a need or problem, to the final satisfaction of the need, or accomplishment of the desired goal or objective. The steps which usually occur in such a social action program can be organized into a logical time sequence model to provide a framework for a planning and analyzing action oriented educational programs.[2]

All educational change, just as other social action, takes place within the contest of social systems. The beginning of all educational change is with convergence of interest and definition of need by a few people. This stage might be called the initiation stage of an action educational program.

In most communities or counties there have been previous experiences with similar educational action programs which should be considered in planning any new educational effort. Seldom are all the social sub-systems of a county involved in any one action program. This being true, it becomes necessary to outline the relevant social systems, or the "target system," for a particular educational effort. Initiation of action of the "sounding board" variety and legitimation, or sanction giving, are important considerations in building the foundation for any action educational program.

The process of organizing and planning for action in such a way that the educational need is defined and becomes the felt need of the people, around which those in the relevant social systems commit themselves to action and set objectives or goals for themselves, is another broad area of concern in the conduct of action educational programs. At this stage, the specific series of learning experiences that will be needed to attain these objectives and goals can be set up in a plan of work. A plan developed in this manner, and including evaluation for determining progress, can then be carried through.

The Extension Teaching or Program Development Cycle

In extension educational program development, five reasonably distinct stages are recognized. Figure 1 and accompanying description of each stage will help clarify this total process and will further emphasize the importance of evaluation throughout each extension teaching or program development cycle.

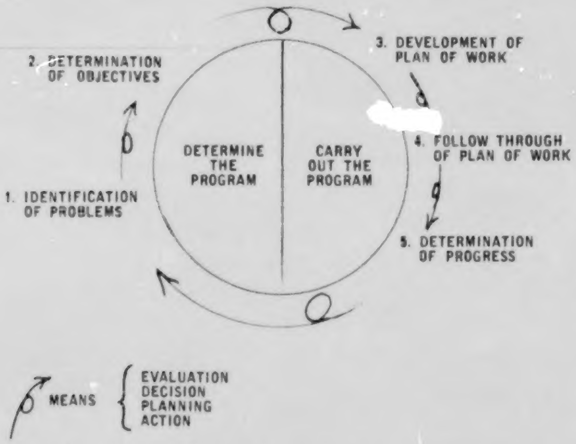

2. DETERMINATION OF OBJECTIVES
3. DEVELOPMENT OF PLAN OF WORK
DETERMINE THE PROGRAM
CARRY OUT THE PROGRAM
4. FOLLOW THROUGH OF PLAN OF WORK
1. IDENTIFICATION OF PROBLEMS
5. DETERMINATION OF PROGRESS
MEANS
EVALUATION
DECISION
PLANNING
ACTION

Figure 1.—The Extension Teaching or Program Development Cycle

Stage 1. **Identification of the Problem**—Through collection of facts and analysis and interpretation of situations by the people working with the extension worker, the significant needs and interests of the people are identified. Those areas of needs and interests wherein it is believed improvement can be accomplished, and which lie within the scope of legislation applicable to the Cooperative Extension Service, then become the foundation for an immediate and/or long-range program. This stage can be evaluated from the standpoint of both content and processes or procedures used.

Stage 2. **Determination of Objectives**—After the needs and interests have been identified, the people, working with the help of extension personnel, can decide on their objectives. This involves a definition of what is to be accomplished in relation to the various needs and interests, both immediate and long-time, as well as how and by whom these accomplishments are to be instrumented. Criteria for this stage can be established and evaluations of both content of the objectives and procedures followed in their determination can be made.

Stage 3. **Development of a Plan of Work**—After the needs and interests are identified and the objectives are agreed upon, the plan to be followed for putting the program into action must be developed. Sound principles of social action must be followed in developing such a plan. Involved are such things as determining, (1) the specific jobs that need to be done, (2) the subject matter that is needed, (3) the teaching techniques to be used, (4) the activities to be undertaken, (5) the division of responsibilities, and (6) the calendar to be followed. Evaluation based on established criteria for plans of work is very helpful at this stage of the program cycle.

Stage 4. **Follow Through on the Plan of Work**—A good plan of work presents the most effective way, considering existing circumstances, to accomplish the agreed upon objectives. An action program, based on such a plan of work, and built on a foundation of sound education and social action principles, requires systematic and persistent effort from all concerned. Evaluations can be made of what is actually done in the action program from standpoints of both content and procedure.

Stage 5. **Determination of Progress**—The extent to which objectives are being accomplished is the basis for determining how well a program has succeeded. Evaluation of progress helps determine what remains to be done. When objectives are satisfactorily reached, new ones can be included in the program and plan of work for the following year. If the objectives are not accomplished satisfactorily, a revised plan of attack on the same problem may have to be carried over into the new program. In addition to data for future or followup programming, progress evaluation also provides tangible and objective data for use in annual and other periodic reports.

Active participation of the people in all five stages is the core of the teaching-learning process in Extension. Each of these stages can provide many experiences for those involved and it is through such experiences that learning takes place.

As indicated, each stage can be evaluated. This is important because the accomplishments and successes attained in any one stage are indicative of how much is likely to be attained in each subsequent stage of the extension teaching cycle.

Summary

In this chapter we have stated that the Cooperative Extension Service was brought into existence in 1914 to perform an educational function for persons not in residence at land-grant institutions in the fields of agriculture and home economics. Through the years it has developed an educational philosophy which calls for reaching and working with the people where they are in matters of educational, occupational and socioeconomic maturity. Emphasis always has been on **how** rather than what to think.

We have said that the teaching-learning process consists of producing desired changes in human behavior and that it is a dynamic or action oriented process. We have said that evaluation is a key concept in the process and have outlined the five stages of the extension teaching cycle with notations as to how evaluation fits into each stage.

Next we will go into greater detail as to what evaluation is and why it is so fundamental to all aspects of Extension.

References

1. Association for Supervision and Curriculum Development. **Guidance in the Curriculum, 1955 Yearbook.** National Education Association, Washington, D. C.
2. Beal, G. M., and Bohlen, J. M. **The Diffusion Process.** Iowa Agr. Ext. Serv. Special Report 18, Ames, 1957.
3. Brunner, E. deS., and Yang, E. H. P. **Rural America and the Extension Service.** New York, Teachers' College, Columbia University, 1949.
4. Cantor, N. **The Teaching-Learning Process.** New York, Dryden Presses, 1953.
5. Deese, J. **The Psychology of Learning.** New York, McGraw-Hill Book Co., Inc., 1952.
6. Hannah, J. A., (Chr.) **Joint Committee Report on Extension Programs, Policies and Goals.** Washington, D. C., U. S. Dept. of Agr. and Assn. of Land-Grant Colleges and Universities, 1948.
7. Kelsey, L. D., and Hearne, C. C. **Cooperative Extension Work, 2nd Ed.** Ithaca, Comstock Publishing Associates, Cornell University Press, 1955.
8. Lancelot, W. H. **Permanent Learning.** New York, John Wiley and Sons, Inc., 1948.
9. Miller, P. A., (Chr.) **The Cooperative Extension Service . . . Today: A Statement of Scope and Responsibility.** Subcommittee on Scope and Responsibility of the 1957 Extension Committee on Organization and Policy.
10. Tyler, R. W. **Basic Principles of Curriculum and Instruction.** Chicago, University of Chicago Press, 1950.
11. Wilson, M. C., and Gallup, G. **Extension Teaching Methods.** U. S. Dept. of Agr., Fed. Ext. Serv. Cir. 495, Washington, D. C., 1955.

Chapter III

THE PLACE OF EVALUATION IN EXTENSION

Joseph L. Matthews

Our objectives in this chapter are to help the reader to: (1) understand the role of evaluation in Extension, (2) know some of the benefits that can be gained through systematic evaluation, (3) know what can be evaluated, and (4) feel that it is important to evaluate Extension work. While so doing, we hope to provide the reader with a usable answer to the question, "What is the place of evaluation in Extension?"

We evaluate so that we can have some measure of the desirable, and undesirable, outcomes of educational action that we take to move a person or group of persons toward desired extension goals or objectives. We are using action here in the sense of attempting to change human behavior in desired ways.

Why Evaluate?

No doubt, this question often has entered your mind. You may already have your answer, but we would like to discuss our's briefly. As we stated in the opening chapter, everyone does a certain amount of evaluation, sometimes consciously but quite often unconsciously. We ask ourselves, "How am I doing?" at practically every stage of our work. Every job we do (writing a news story, giving a talk, making a farm or home visit) brings up this question.

We may not frame the words mentally, but after we have done something we may feel good about it, or we may be only moderately satisfied with it, or our reaction may be one of dissatisfaction with what we did. By these reactions we have passed judgment on our activity and have found it good or bad. Making judgments in this manner about the value of what we have done is a kind of evaluation.

On a slightly higher plane is the kind of evaluation we do by assessing what others say or do about our work. If a person who has attended one of your meetings tells you that he or she enjoyed the meeting, you feel good. If several persons tell you the meeting was a good one, you feel even better about it and think that it must have been a pretty fine meeting.

Expressions like these have some value as indicators of success. When nothing better can be done about evaluating our work, we come to depend a lot on this kind of evidence. However, we must not depend entirely on such expressions from people who participate in our extension activities. Quite often, persons making judgments of this sort have no real understanding of what the educational purposes of Extension are. For this reason, they have no sound basis for making judgments about results.

Sound evaluation always must be in terms of what we set out to do—the objectives we have in mind. Therefore, expressions of the people involved, as well as all other measures, must be appraised in the light of what we expected to accomplish when we planned the educational activity.[2]

In addition to the danger that what people tell us may not be relevant to our purposes, there are other reasons why such evaluations must be supplemented by evidence from other sources. Persons who have participated in our program and who are our friends are reluctant to be critical even when asked to be. This results from not wanting to hurt our feelings and from an inability to separate their feelings toward us from their objective appraisal of our professional activities. Later in this manual we will have more to say about the complicating factors of the human element in scientific evaluation.

The preceding paragraphs lead logically to this statement: "We all evaluate our work in some way so why not do it well?" If the first part of the statement is not true, then experience would have no value and there would be no improvement on the job as a result of experience. Doing evaluation well requires considerable self-discipline in applying scientific methods, collecting and analyzing facts, and making judgments about their implications. This manual is all about doing evaluation well by applying scientific methods to evaluation.

What Can We Gain from Evaluation?

The list of benefits we can gain from evaluation could include many more than the five important ones that we will mention here. First, extension workers have to make endless decisions and then act according to what they understand to be the mandates of their decisions. For example, a county extension worker may believe that appointing the chairman of an action committee will result in better committee work than when the committee members choose their own chairman. Much later, and probably by accident, it may be discovered that the expected result actually did not occur.

The longer a practice has been followed, the harder it is to be objective about its limitations, and the harder it is to get at making needed changes. One reason for this is a predisposition to see the expected results even though they are not actually in evidence. Preconceptions about results often make perception of actual happenings so selective that unexpected or undesired ones go unnoticed. Evaluation is the method of trying consciously and objectively to find out whether or not certain activities actually did lead to the results that were anticipated.[1]

Good evaluation is essential when it comes to making periodic checks on the effectiveness of educational programs to decide where improvements are needed. These checks serve a function similar to the balance sheet in a business enterprise. When the sheet reports results that are less than expected, owners or managers of the business look over the units which are not doing well to see where changes are in order. In either case, decisions so basic to successful functioning of the organization must be founded on accurate measurements and observations, or changes may not be made in the best direction. Surely we all agree that change for change's sake is not good long-run action.

A third benefit of good evaluation is the feeling of security it can give the worker who is doing a good job. Extension workers can become so involved in their work that it is easy for them to worry about their effectiveness and to have doubts as to whether they are really accomplishing their objectives. This concern is all right if it leads to careful appraisal and constructive measures to improve programs. However, when systematic evaluations are not made, the tendency is for staff members to feel even less secure and to retreat to activities that have surface indications of tangible results, although they may be less important to the work than others which they abandon.

We may seek security by emphasizing extension activities to which people will respond in large numbers rather than others that could result in basic educational changes in the long run merely because we need the immediate assurance of favorable popular reaction. Scientifically done evaluation will give equally convincing assurance when good work is being done and we know the criteria will be valid extension objectives rather than popular appeal.

A fourth benefit lies in the area of public relations. There is no better basis for sound public relations than a careful and comprehensive evaluation. Rather than superficial impressions or value-clouded opinions, it can provide concrete evidence that can be adapted readily to the needs of people working in the various communication media.

These folks are happy to help inform the community of the contribution of your educational program when they see the facts. They do not have the feeling they are "going out on a limb" for you when they are handed the ammunition to deal with criticism. They know that critics can be converted to constructive cooperators when concrete evidence of accomplishment can be produced.

A final benefit we will discuss is the matter of what evaluation does to the professional attitude of the worker who uses scientific evaluation techniques. In the process of stating his objectives in systematic fashion, specifying the desired changes, analyzing the teaching activities designed to accomplish the objectives, and appraising the results and probable need for further action, he is very likely to adopt a more scientific approach in future undertakings. Inevitably, he will adopt more objective and systematic ways of thinking about his work and his accomplishments.

Stages of Evaluation

We can think of evaluation as having two main stages. Early in the program we can study arrangements and procedures for doing Extension work. Suppose we refer to this as **means evaluation.** Evaluation of means is especially useful in the early stages of an activity or program before it has been going long enough to accomplish the desired educational results. We can study methods, or means, and decide whether they are worthy of the program in which they are being used.

Examples of means evaluation would include appraisal of the organization of people for extension committees, clubs, special interest or commodity groups and the like. Educational materials used such as bulletins, leaflets, newspaper articles, written programs, and plans of work are included among the means which should be evaluated. The procedures for planning programs and conducting club meetings also should be considered.

Evaluation in terms of means alone is hardly adequate, however, and should be thought of as part of the evaluation sequence rather than the whole of evaluation. Insofar as the extension program is concerned, the end products of extension work, the educational changes that result from extension teaching, are what must eventually be evaluated. We must make **ends** evaluation. We need to assess the changes in knowledge, understanding, attitudes, skills and abilities of the people who are affected by Extension educational activities.

Purposes of Evaluation

In addition to kinds of evaluation in terms of stages of the program, we can think of evaluation in terms of objectives. We can think of it in terms of pure fact finding or **measurement;** or we can think of it as a basis for **interpretation.** In the first is involved matters of assembling measurable data on changes presumed to result from the teaching device—observable behavior changes as well as changes in material things. In the second we make considered judgments about the data that were collected by means of the measuring techniques.

Although most of us can recall when we stopped at the first stage, or when we observed others stop there, we should have no difficulty in acknowledging that the evaluation mission is not accomplished unless we face up to the interpretation stage. At some point judgment must be made about the degree of success or lack of success that was experienced in the program.

Who Should Do Evaluation?

Evaluation, insofar as possible, should be done by the extension workers who are actually responsible for the program that is being evaluated. It improves their understanding of the program and their work in general. In using the methods of science to evaluate their work, they accumulate reliable evidence which helps them define future problems more sharply. They draw on these experiences and findings for action hypotheses that give promise of helping to reduce or eliminate difficulties in their work. They test out promising procedures on the job and accumulate evidence of their effectiveness in the next evaluation sequence. Eventually, after obtaining sufficient evidence of effectiveness, they feel justified in generalizing within limits and can turn to other objectives.

An important psychological consideration in evaluation is the fact that the people who stand the best chance of benefitting professionally from the experience of the evaluation are the ones who do the evaluation. They can see at first hand the effects of the methods and procedures and can get speedy answers as to how good new and promising ways of doing the job really are. For this reason persons responsible for the program should make every effort to do the evaluation or be a part of the evaluation team.[1]

Another consideration in evaluation is the problem of communicating results to the ones who can use them in their work. The communication problem is solved in proportion to the degree of participation by the people who are responsible for the program. When the responsible persons are fully involved, or when they take total responsibility for evaluation, the communication problem virtually is eliminated.

Sometimes, however, the complexity of the problem being studied or the technical competence needed, may make it impossible or impractical for the extension worker to conduct the evaluation effort independently. Nevertheless, drawing upon the skills of others for assistance in evaluation need not impair communication nor minimize the value of the experience to the extension worker whose work is being studied. Experience has shown that there are some things that extension workers have neither the time nor the technical competence to handle by themselves. These, very properly, are left for State extension research persons and for the resident teaching or experiment station research staffs to deal with.[4]

Summary

In this chapter we have attempted to impress our readers with the importance of evaluating our work on the most scientific level the particular activity warrants. We have listed and elaborated on five important benefits to be gained by evaluating. We have mentioned that there is both means and ends evaluation, and that both have a place in extension teaching, but that we must not stop after doing the former when it is possible to do the latter as well. We have stated also that evaluation can be a matter of measurement alone, or it can be carried to its logical conclusion by interpreting results of the measurements to determine effectiveness of the action that was evaluated.

Closing remarks were devoted to explaining why extension persons responsible for the program either should be the ones to evaluate it, or should be as closely involved as possible in the process. At the same time, they should feel free to call on professionally qualified counsel and leadership in situations calling for more facilities than are at their own disposal.

In the next chapter we will spend some time discussing what is involved in scientific methodology and what is necessary to consider before doing a scientific evaluation.

References

1. Corey, S. M. **Action Research to Improve School Practices.** New York, Columbia University Press, 1953.
2. Matthews, J. L. **How Are You Doing—In Farm and Home Development.** Talk given at the Southern Regional Farm and Home Development Workshop, Rock Eagle, Georgia, September 21, 1955.
3. Smith, E. R., Tyler, R. W., and evaluation staff. **Appraising and Recording Student Progress.** New York, Harper and Brothers, 1942.
4. Tyler, R. W. **Organizing for . . . Using . . . Research in Extension: Summary of an Extension Research Workshop.** May 6-10, 1957, Washington, D. C., Federal Extension Service, U. S. Department of Agriculture, 1957.

Chapter IV

THE SCIENTIFIC APPROACH AND EXTENSION EVALUATION

Fred P. Frutchey

People are curious. They have always been curious. They want to find out what makes things "tick." They want to understand themselves, other persons and the world in which they live. They try to study and investigate puzzling questions.

This curiosity is not confined to adults. At a very early age people are curious. The "why" questions of children are well-known to parents and teachers.

People are not only curious but they also have problems and must make decisions. Very early in the history of man, the problem of measuring time was recognized. Man put his ingenuity to work to find a solution. The measurement of time progressed over the years from such gross units as three moons ago, through the hour-glass stage to electric clocks and self-winding watches.

How do people find out things? How do they make up their minds? How do they know what to believe, and what to do?

During the history of man's development, people have used a variety of approaches to gain understanding and to solve problems. They have used custom and tradition, authority, personal experience and intuition, and the scientific method.

Custom and Tradition

We use custom and tradition many times in our lives in reaching decisions—and with no harmful affects, as in buying clothes, and many times for highly essential and valuable purposes, as for example, courtesy and manners.

But when we use custom and tradition for **all** our decisions, improvement and progress "go out the window." If farming decisions were dictated entirely by the farm practices of our ancestors, extension agents would have a difficult time implementing the findings of the experiment station.

We would have no automobile, no radio, no electric power, no telephone and a host of other things because they were not our custom or tradition. The past would be perfectly preserved.

What would happen if extension agents made all **their** decisions on the basis of custom and tradition?

Authority

Custom and tradition did not always provide a satisfactory answer to curious people; nor solutions to urgent problems. People appealed to "authorities," those in whom they had confidence and whom they thought should "know."

They went to the "wise men of the tribe," to the oracles, to the "medicine men" and up through the ages to the authorities of today. Much of our teaching in schools and colleges is accepted on authority. We cannot possibly explore and comprehend all knowledge. So we depend on authority.

People frequently consult technical and professional authorities like physicians, county extension agents, teachers, lawyers, bankers, and so on.

Football players are governed by the rules of the game. Referees make their decisions on the basis of the rules. Card players are governed by the rules of the game. Extension workers are governed by the policies of the institution.

But authority must get its knowledge in some way and arrive at the rules of the game, the policies of an institution, or the advice of experts.

Personal Experience and Intuition

Personal experience is a widely used basis for making decisions. Notice how often it is resorted to in argumentative situations, and how it is given as a reason for our decisions. This is done consciously.

But intuition is somewhat different. Intuitive insights seem to come from nowhere. They are hunches. Have you ever worked long and hard on a problem and have finally gone to bed without solving it? Upon awakening in the morning, the solution came to you like a flash. Where it came from is difficult to say. It just came.

Or have you ever had to make a judgment when you know what it should be but could give no good reason for it? These are two examples of intuition. Such decisions or judgments seem to be based on personal experience and to have come out of the subconscious.

Long practical experience in a field of work tends to develop this intuitive power in us and to make it easy for us to use intuitive judgments. They are frequently right, and they are frequently wrong.

Personal experience can be very misleading at times. For example, we may draw a conclusion from a single instance. Note in Chapter I other common errors we frequently make during personal experience in getting information and forming judgments.

Following the rules of evaluation will help us to use personal experience as a more satisfactory basis for making decisions and checking on our intuitive judgments.

Scientific Method

Toward the end of the 16th century another basis for making decisions was tried. This was checking to see, testing it, experimenting with it, trying it out.

The often quoted illustration of the use of the scientific method was Galileo's experiment. In Italy in 1589 Galileo and others were arguing about whether a heavier weight falls faster than a lighter weight. Some said that it stands to reason that a heavier weight will fall faster.

Galileo proposed that they drop two weights, one heavier than the other, from the leaning tower of Pisa at the same time and see which one hit the ground first. This was a most unusual way of deciding a question at that time.

Eventually Galileo tried this way. Both weights when dropped at the same time reached the ground at the same time. They dropped with equal speed, neglecting air resistance. What an odd way to "prove" anything! This was a revolution in thinking. It opened the way for discovery. It freed the minds of people to find out new things.

Today we take Galileo's method as an acceptable method—or do we? How many times do we make decisions, form judgments, jump to conclusions without checking to obtain a better basis for our actions?

The method used by Galileo is the method which has made it easier to understand ourselves and the world in which we live. It is the method which has made possible the progress of today. It is the method of evaluation.

Characteristics of the Scientific Method

There are certain characteristics which identify the scientific method of making decisions, forming judgments or drawing conclusions. We will discuss five of them briefly at this point and elaborate on them in succeeding chapters.

1. **Factual**—The scientific method is a way of working by which observable and demonstrable evidences can be obtained and upon which people can agree. If we want to find out whether farmers are adopting certain practices or have certain attitudes, we can do as Galileo did—look and see.

2. **Analytical**—Most of life is very complex and involved. It must be broken down into its component parts so that the human mind can deal with it. Most processes in Extension are too complex to study as entireties.

Even leadership, a component part of extension education, is not manageable until it is analyzed to the point where there are "handles" that we can take hold of and understand. The big intangible problem must be broken up into small more tangible problems.

Evaluation may seem to be a very complex process. Through analysis it can be broken down into its component and understandable parts. Later in this chapter steps in evaluation are presented and succeeding chapters amplify these steps.

3. **Impartial**—Scientific **method** is a **method** of attacking problems. Scientific **attitude** is how we **feel** about using scientific method. It is a disposition or willingness to apply scientific thinking.

A debater argues to prove his side of an issue. He canot have a scientific attitude. If he did he might weaken his argument, and lose his case. He dare not present both sides of the issue. The opposing debater presents the other side.

A lawyer, like a debater, cannot argue for both sides. He supports and strengthens his own side of the case only.

A judge or jury, on the other hand, must have a scientific attitude. They must look at both sides and weigh the pros and cons of evidence impartially.

A scientific attitude is a willingness to look at our interests, prejudices, and loyalties. It is a willingness to examine and re-examine our beliefs and explore new ideas.

A scientific mind seeks what is just, true and beautiful. It is alert, curious, interested and a little excited about the world. It is an open mind (not to be confused with an empty mind). It is an attitude of mind that is learned, not inherited.

4. **Reliable**—Have you ever noticed that just before Christmas children are on their good behavior? During the rest of the year they are just normal children. If you observed those children **only** just before Christmas or **only** at some other time of the year when their behavior was at low ebb, you would come to opposite conclusions.

Reliability means that we should have a sample of representative observations and be careful about drawing conclusions from a single observation at any particular time.

5. **Objective**—Each of us uses our subjective judgment in interpreting what we see and in draw-

ing conclusions about our observations. Do others see what we see and do they draw the same conclusions? The degree to which they do so is objectivity.

If we look at a dog, each of us would agree that it is a dog. Except perhaps at a dog show, we might see an entry and ask, "Is **that** a dog?"

Objectivity is the agreement among people of what they observe and of their interpretations. In evaluation we are concerned about agreement on the evidence we are looking for and upon recognizing it when we find it.

We are concerned about agreement and understanding of the methods used in evaluation. We are concerned about agreement on the interpretation and conclusions reached.

Many arguments result from differences of opinion. The use of scientific thinking helps to dispel these heated arguments. When people clarify what they are arguing about, they often find that they are not arguing about the same thing.

Have you ever been in a discussion where after an hour the group found that all were not talking about the same thing? Clarifying the question makes for better communications, more efficient use of time and better human relations.

The above examples illustrate the need of objectivity to assure understanding of the question at issue.

Objectivity is increased by applying numbers to our observations. Whereas "a few persons" may mean 3 or 10 persons, "5 persons" is very definite. "Five persons" is more objective than "a few persons." All of us can look and see if there are **five** persons.

At one time length or distance was not expressed quantitatively. Such expressions as "a piece way down the road" or "quite a long way from here" were used. Now we measure distance in the number of inches, feet, yards, miles and even light-years.

The thermometer is a more objective measure of heat than our feelings of being hot or cold. It has been said that progress in scientific study has accelerated with the development of satisfactory quantitative measurements. These are used extensively at the upper end of the scale of degrees of evaluation.

Many things in the behavioral sciences today, however, cannot be quantitatively measured; they can only be described. As time goes on, and better measuring devices are built, more quantitative measurements will be made.

Science and Scientific Method

To some people research is associated with subject-matter fields. Science is considered to be in the fields of chemistry and physics. To these persons the biological fields such as medicine, plants, animals and food are considered on the fringe of science. They believe the social studies like education, psychology, sociology, anthropology, and similar fields are clearly outside the realm of science.

Other persons claim that all of these fields are in the realm of science. The difference of opinion raises the question, "What is science?" Is science a field of subject matter or is it a method of thinking? To the latter group of people, physical, biological, or social phenomena are not in themselves science until the human mind brings to bear on these phenomena its powers of observation, analysis, synthesis, and verification in organizing these phenomena for use by people.

In this sense, science as a method of thought has application to all fields of subject matter. Scientific method involves the logical thought processes based on the recognition of assumptions and definitions. It involves the collection and interpretation of data and the control of factors which might influence the result of an experiment, a study or an investigation.

Mathematics can be applied to all the fields, more accurately to some than to others. In similar manner experimental controls can be applied to all the fields, with much more difficulty in some than in others.

The past 50 years have shown marked progress in the application of scientific method to problems in the field of the social sciences. Evaluation in human behavior has shared in this progress. Today we have evaluation in extension work as an application of scientific method with various degrees of rigor depending upon practical conditions and usefulness of the results.

Five Questions

We have found that by asking ourselves five simple questions we can get off to a fast start on a problem requiring evaluation. Slight modifications will make them fit most situations.

1. **What Do You Want to Know?**—The first thing you need to do is decide exactly what you want to find out. What are the questions you want answered by your evaluation? Work out these questions as carefully as you can, because they determine and guide the conduct of the evaluation, and because the answers will help you in your work.

The questions will guide you in stating the specific questions for the questionnaire or other devices you use. Analysis can be made of the replies you receive to these questions. The report of your evaluation can be organized around these questions. In fact the report will be the answers to these questions.

By carefully determining what you want to know, an initial, vague, general feeling of concern can be broken down into questions which you can tackle.

2. **Where, and How Will You Get the Information?**—This step has two parts; where to get the information, and how to get it.

Where means from whom or from what other source. Will it be from adults, from men, from all farmers in a county, from dairy farmers, or from Grade A milk dairy farmers? Will it be from all persons in any one of the groups or from a sample of persons in one of the groups?

Other sources of information are records and reports, such as annual reports, the census, written programs, minutes of meetings, and so on.

How you will get the information refers to mail questionnaires, observations, interviews, tests, check lists, score cards, case studies, and so on. Forms are prepared on which the information can be recorded. They are used to facilitate handling the data and so you do not have to depend upon memory.

When you have thought through and know the information you want and also where and how you will get it, you are well on your way in planning an evaluation.

3. **Who Will Collect the Information?**—The next step is to decide, whether you or a group of other persons will collect the information. What training will be necessary if you have others collect the information?

4. **How Will the Information Be Analyzed?**—Now is the time to go back again to the first step—the information you wanted. The replies on the record forms are organized and summarized to answer the questions formulated in step one.

5. **What Does It Mean?**—What are applications of the results to extension work? In what way can they be used to make our work better? After knowing these results, what changes should one make in his way of doing extension work?

Often the changes are not big revolutionary changes. Many times there are little ways, such as doing something different in organizing a group, or in writing a letter, or in a farm or home visit. You may get to know your group better and understand why some do not attend meetings or why they are not interested in new practices.

All of these steps are thoroughly discussed in succeeding chapters.

An Evaluation Outline

The five questions we have been discussing can be expanded into an evaluation outline as a plan of action that you can use in making "do-it-yourself evaluations" and "extension studies."

Although the steps have a certain element of sequence to them, it is wise to keep all steps pretty well in mind while working out the details of any one of them. All the steps are not usually taken in any given order. Never consider only one step at a time without thinking about the other steps, since all steps are related to each other.

1. Need for the Evaluation
   a. What extension project, problem, activity, job, method or situation do you want to evaluate?
   b. Why evaluate it?
   c. How can the results of the evaluation be used in your extension work?
2. Purposes of the Evaluation
   a. Is the evaluation,
      (1) an analysis of a situation, or
      (2) an evaluation of teaching objectives?
   b. What questions should be answered by the evaluation?
3. Questions to be Answered by the Evaluation
   a. If analysis of a situation, clarify the kinds of information needed to answer the questions.
   b. If an evaluation of teaching, clarify the teaching objectives. Analyze the teaching plan.
      (1) Review what has been taught.
      (2) How it was taught, and to whom.
      (3) Changes in behavior expected as a result of the teaching.
4. Sources of the Information
   a. People—farmers, homemakers, local leaders, club members, non-members, extension agents, etc.
      (1) Do you need a sample of these people?
      (2) How will you draw a sample?
   b. Recorded information — reports, census, etc.
5. Collecting the Information
   a. How? Interviews, mailed questionnaires, observations, etc.
   b. By whom? Person making the evaluation, extension personnel, local leaders, etc.
6. Selecting or Constructing a Record Form
   a. Kind of form; questionnaires, interview forms, tests, observation sheets, rating scales, check lists, score cards, anecdotal records, etc., to be used.

   b. Data about
      (1) the situation to be studied,
      (2) evidence of progress toward the teaching objectives, or
      (3) "face data" to be collected.
   c. Formulation of questions or statements
   d. Physical set-up of the record form
   e. Pretesting and revisions

7. Analysis and Tabulation of the Data for Use
   a. Classification and sorting of data to answer questions in Step 2 b.
   b. Work tables needed
   c. Coding of free-answer questions
   d. Method of tabulation
   e. Sorts and subsorts to bring out relationships

8. Interpreting, Reporting and Applying the Findings
   a. Preparation of tables, charts, and graphs
   b. Summary of findings
   c. Applications to extension work.

Summary

The curiosity and necessity of people to learn about themselves and their environment has occasioned them to try various methods of getting information and making decisions.

The most satisfactory method has been the scientific method. It is the method of careful observation and control of subjective bias. Scientific method can be used in solving all kinds of problems and is not limited to the physical sciences.

Five simple questions in evaluation are:

1. What information do you want to get?
2. Where, and how will you get the information?
3. Who will collect the information?
4. How will it be analyzed?
5. What does it mean?

These questions were expanded into an evaluation outline.

The scientific method is a method of thinking and we can apply it in our everyday work.

The four chapters of this introductory section are designed to explain evaluation, its relation to extension education, how it is used in extension work and the scientific approach to the study of problems.

In Section II, methods used in the evaluation process will be discussed in detail.

Reference

1. Rummell, J. F. **An Introduction to Research Procedures in Education.** New York, Harper & Bros., 1958. Chapter 1.

THE EVALUATION PROCESS

Chapter V

GOALS AND OBJECTIVES

J. Neil Raudabaugh

In a more comprehensive text we might be justified in devoting a chapter to outlining the shades of difference among concepts such as purpose, objective, intention, goal, need, aim, drive, want, and others. For some purposes, we in Extension do make certain distinctions among them. For immediate purposes, however, we will use them interchangeably, but will tend to use "goal" or "objective" most frequently.

Need for Objectives

How often have you heard remarks like this? "Everyone knows what we are trying to accomplish; let's get on with the job." Statements of this kind are not uncommon among people working with Extension educational programs. Very often the importance of spending the time necessary to determine and define goals is not realized by laymen, and all too often professionals who work regularly in the field of extension education make the same error. Because of this, we may find professional educators, lay leaders and educational groups so concerned with their on-going procedures and activities that they lose sight of what it is they are trying to accomplish. They fail to realize that educational objectives are the criteria by which content is outlined, materials are selected, teaching procedures and learning experiences are developed, and progress toward accomplishment — and accomplishment—are evaluated.

Kinds of Objectives

In considering goals or objectives in Extension, it is essential that we think of them in relation to the people involved, or to be involved. Some objectives will be **group** objectives and some will be **individual** or family objectives. Once goals and objectives are determined, they exert a definite influence on the activities and behavior of the people participating, as well as on the extension worker involved. Time spent by extension agents helping people to define their goals and objectives can do much to help them gain that needed direction and purpose which will, in turn, motivate them to increased achievements and satisfactions.

Goals and objectives indicate what people are trying to accomplish; they identify the target toward which the action program is oriented. They can be long-time or immediate, they can be general or specific. In an educational program it is desirable to distinguish between long-time (ultimate) and short-run (immediate) objectives. Both have their place in extension teaching, but it is important to the total program that they be compatible.

In any consideration of objectives it is also essential to distinguish between **broad** objectives and **specific** objectives. The latter are more apt to be achievable, dynamic and capable of doing much to help promote action and participation. The former are achieved with greater difficulty, largely because progress is not so apparent because of the time span, and because agreement on feasibility is not so universal.

People working with action educational programs sometimes adopt a set of abstract objectives and then proceed to consider them as though they are short-run and easily achievable. Such objectives especially for extension educational work, fail to provide needed direction. They limit prospects for achievement and ultimate success of a program, so necessary to stimulate in the people concerned the desire to strive for growth and development.

People, both individually and in groups, respond best to specific objectives that they believe are achievable. If educational programs are intended to stimulate people to participate, dynamic and achievable objectives are requisites. Such objectives also serve well as the basis for identifying evidence of accomplishment for purposes of evaluation.

Levels of Objectives

The relationships and differences between objectives of extension workers (teachers) and those of their clients (learners) are not always realized. This failure results in much ineffective teaching, and contributes to the development of unfavorable attitudes toward educational programs.

Extension workers, who are not sure what the people's objectives really are, have to conduct an educational program they believe is needed, present subject matter they consider appropriate for the program, and comfort themselves with the hope they are achieving desired objectives. Actually, they may be producing few, if any, changes in the people they are attempting to help. In cases of this kind, the people themselves, while failing to understand and accept the objectives suggested by the extension worker, may have no idea of their own objectives.

When dealing with this frequent conflict in objectives, it has been found helpful to think of objectives on various levels, each with a place and function. Burton, in his book, "The Guidance of Learning Activities," has classified objectives into four levels.[3]

1. **Objectives of Society**—The general aim of every society is to secure the "good" life for all of its members. Surely all of us have spoken glibly at one time or another of high standard of living, maximum production, favorable balance of trade, religious freedom, and such generalities. Remote and abstract objectives of this kind have a place in designing all-inclusive ultimate ends, but they are of little value as objectives for an action oriented extension educational program.

2. **Objectives of a Specific Organization or Group**—This is the level of objectives such as are stated in the printed materials relative to State and federal activities of the Cooperative Extension Service. The objective toward which these agencies are directed is improvement of the economic, social and spiritual well being of farm families in general. Objectives at this level have the large group rather than individuals or families in mind and, for this reason, are not suitable for designing action extension educational programs below the State level.

3. **Objectives of the Extension Agent**—An example of an objective at this level might be to develop in livestock farmers an understanding of the merits of stilbesterol in livestock feeding. A broader one at the same level would be to interest parents of junior and senior high school children in ways of improving their understanding of teenagers.

Characteristics of objectives at this level are that they consist of what changes agents desire to bring about in the people with whom they work. They serve as evidence of the ability of the agent to translate extension objectives into action objectives. However, agents who work exclusively with objectives appropriate to this level can never be positive how well their plans fit the needs of their particular clients.

4. **Objectives of the Clientele** — These are the things the people wish to do or accomplish. A farmer may desire to increase his income by ten percent by improving his cattle feeding operations; a homemaker may want to make a coat or dress in order to reduce the social stigma of an inadequate wardrobe and still keep within the family budget.

An extension worker cannot always be aware of people's purposes or objectives. He may have to set out deliberately to learn what they are. However, an alert worker can spot some of them as they emerge in the course of routine activities. He must think through situations and attempt to foresee what specific desires or needs are likely to reach the awareness level of the people. Often it is necessary, and desirable, for him to do a little stage setting and manipulate the environment to help the needs and purposes emerge to be recognized.

These four levels of objectives are not related to each other in a strictly additive manner, but they cannot be entirely incompatible with each other. The objectives of any teaching-learning situation in Extension must originate in the situation, itself. They should be derived from factors present in the situation which are of concern to the people.

The immediate objectives of the people are the progress objectives. These progress objectives bear upon their more remote aims. The extension agent, as teacher and educator, must exercise leadership and guidance to see that a balance in emphasis between immediate and ultimate objectives is achieved.

Those most directly concerned often do not realize the need for this balance. Agent attempts at developing awareness of this need can be confusing, and have been known to antagonize the clientele. Objectives at any level need to be consistent and compatible with those at other levels, just as immediate or short-run objectives need to be compatible with ultimate or long-time objectives.

Objectives of the extension worker and his clients need not be similar in form but they need to have a common core. The people's objectives are those which they believe can be accomplished through activities in problem solving projects of their own design. The extension worker's objectives are the desired educational outcomes which he hopes the people will achieve.

The extension worker's role is to organize and guide learning experiences so that desired educational results (level 3 objectives) can be achieved while the people are achieving their level 4 objectives. Failure to recognize the difference between these two levels of objectives has caused much ineffective teaching. It contributes to the formation of detrimental attitudes and to reduced participation by the very people who should be most concerned.

Stating Educational Objectives

Along with an understanding of levels of objectives, there is the problem of stating them in a form which will be most useful in developing and evaluating the educational program designed to deal with them. This is an important problem for, unless objectives are stated so that they are useful, they may be filed away unused, and the educational program will not reflect the effort that went into their design. Well stated objectives have two paramount uses. The first is to facilitate the selection and organization of learning experiences; the second is to give explicit direction to the evaluation of educational program efforts.

One rather common practice is to state objectives as activities or things to be done which the extension agents, or others responsible for the program, plan to do for people. Examples of this might be to show how to refinish a piece of furniture, to show how to use insulation materials properly, or to present the agricultural adjustment problem.

Such statements indicate what the person presenting the program plans to do, but they do not say who is to be involved or what is to be accomplished by those involved. Statements in terms of activities to be carried on fail to indicate the kind of change or accomplishment that is expected. Activities are not objectives; they are merely the means designed to accomplish objectives.

Sometimes objectives are stated in the form of general content areas to be dealt with in an educational program. Statements of this kind do not specify what is expected of the people involved. For a group concerned with health, the objectives may be stated by listing headings such as: Sanitation, vaccination, health insurance, and the like. Unless they follow in subheadings, such a listing does not specify the desired changes to be made by the people.

If, as we stated in Chapter II, education and learning are viewed as processes for changing behavior, then objectives may be thought of as the desired changes. Behavioral changes always occur in relation to some particular situation and in reference to some particular person or group. Statements of objectives which give clearest direction to the extension worker are statements which specify the behavior desired, the content in which the behavior is believed to be appropriate, and the people or groups who are expected to change. The following objectives at the extension worker level include the three elements necessary in statements of educational objectives.

People Concerned	Kind of Behavior Change Desired	Problem Area
Members of the homemaker clubs	To increase understanding of	The proposed community center
Consumers of retail foods in the Atlantic Trade Area	To increase knowledge of	The sources of food marketing information
Farm and home development families	To develop interest in and understanding of	How to make a farm and home plan
Farmers in the Little Sioux Watershed	To develop the ability to	Establish contour lines

When objectives are stated so that these elements are present, it is relatively easy to determine what kind of learning experiences should be planned, and what the content of these experiences should be. This degree of explicitness in stating objectives is also adequate to give direction to action for appraising and evaluating results.

The kinds of experiences people need to gain understanding or to develop skills are different from those required to develop mere awareness or interest. Therefore, determination and statement of objectives are important tasks that require the broadest possible insight into the processes of education, the psychology of learning, and the sociology of groups.

Criteria for Judging Extension Educational Objectives

We stated earlier in this chapter that objectives are direction givers for action oriented educational programs, and for evaluation of such programs. If objectives are to serve these two important purposes, it is desirable to establish standards which may be used to judge their usefulness. Here is a series of questions which an extension worker, a committee, or any individual or group might use in judging educational objectives. Ideally, a review of objectives for an extension educational program should result in positive answers to all nine of the following questions.

1. Are the objectives explicit in specifying the area in which the changed behavior is to operate?
2. Are the objectives definitive with respect to kind of behavioral change to be accomplished?
3. Are the objectives stated in terms which identify those who are to be involved?
4. Are the objectives the result of cooperative action by extension personnel and others concerned to analyze the situation and identify the problems?
5. Are the objectives compatible with the general aims of both the Extension Service and the people concerned?
6. Are the objectives specific enough to serve as a base for planning, conducting and evaluating an action educational program?
7. Are the objectives sufficiently limited in number to avoid undue confusion and diffusion of effort on the part of extension personnel and/or others involved?
8. Are the objectives achievable, considering the level of concern, the maturity of persons involved, and the resources available?

9. Are the objectives such that they can relate intimately to both immediate and long-time educational goals, and lead to even higher levels of achievement?

Steps in Setting Goals and Objectives

Extension workers and those with whom they work often do not take time to define educational goals and objectives. Sometimes there is resistance to taking time to define them. This is reflected in the thinking implied by such statements as, "We don't need to waste time doing that," or "Let's get on with the job." Another very common feeling about setting objectives is expressed this way, "Well, everyone knows what it is we are trying to do."

One important step to insure the setting of adequate educational objectives, and their ultimate acceptance, is to provide time and opportunity for the people who are concerned to participate in their determination. Goals and objectives provide little motivation for people who do not take part in their determination.

Another very important step is to explore the ideas and suggestions of those who identify themselves with the particular need or problem. This provides an opportunity for interested people to explore alternatives which might not have occurred to them as individuals, and to secure some sort of consensus. After this is accomplished, statements need to be sharpened up so that they set forth explicitly the three necessary dimensions of well-stated objectives, as outlined earlier in this chapter. These statements then can be tested against appropriate criteria for judging objectives. Once goals and objectives are set, there should be a final check to be sure they are acceptable to and understood similarly by all concerned.

Objectives that have been determined and accepted in this manner meet the criteria for worthy extension educational program objectives and will serve as a firm base for both measurement and interpretation types of evaluation.

Summary

We start this section with a discussion of objectives and how they fit into the scheme of extension teaching and evaluation. By taking the time to state them, we crystallize our thinking on the problem at hand and have a better idea of how to proceed. In so doing we assemble our forces in such a manner that all elements are directed toward a common end.

We mentioned that there are group and individual objectives, that they can be broad or specific, and that they can be immediate or ultimate. We also classified them for use in Extension as objectives of society in general, of the specific organization or group, of the extension agent, or of the people (clientele).

Next we discussed how objectives can be stated so that they are limited in effectiveness by neglecting one or more of the requirements of a well stated objective: Who is concerned? What kind of behavior change is desired? What is the problem area? We also included nine test questions to help determine how well stated an objective is in terms of extension educational goals.

References

1. Association for Supervision and Curriculum Development, **Guidance in the Curriculum;** National Education Association, 1955 Yearbook, Washington, D. C.
2. Brown, I. S. "Working Toward Goals," **Adult Leadership,** Vol. 4, No. 4, p. 4-15. September 1952.
3. Burton, W. H. **The Guidance of Learning Activities.** New York, D. Appleton-Century Co., 1944.
4. **Joint Committee Report on Extension Programs, Policies, and Goals.** USDA and Association of Land-Grant Colleges and Universities, Washington, D. C., 1948.
5. Lancelot, W. H. **Permanent Learning.** New York, John Wiley and Sons, Inc., 1944.
6. Tyler, R. W. **Basic Principles of Curriculum and Instruction.** Chicago, University of Chicago Press, 1950.

Chapter VI

THE PLACE OF TEACHING PLAN ANALYSIS IN THE EVALUATION PROCESS

Mrs. Laurel K. Sabrosky

We cannot find the worth of (evaluate) any activity or program, or estimate the probable worth of a new activity or program, unless we know the educational and organizational objectives of the program. It is just as clearly a fact that we cannot evaluate adequately without a clear picture of the program that was followed in helping the people to reach those objectives, or of a contemplated program that will be followed. This is a basic and vital step in the evaluation process.

To extension workers, evaluation is determining how effective their work is, or how suitable their plans of work are for doing effective extension work. Therefore, we in Extension must find out if our work has contributed, or will contribute, to the accomplishment of desired results.

This is the stage in the evaluation process where we study the methods used, or to be used; the subject matter taught, or to be taught; and the audience reached, or to be reached. This kind of analysis does not answer the question, "Was I effective?" Instead, it helps answer the question, "Have I arranged for a plan that has promise of resulting in an effective program?"

At this stage, we determine what the objectives of the teaching plan were really aimed at during the time the project was being carried. Because extension evaluation is finding the worth of extension work, or assessing the potential of contemplated extension work or some part of it, we cannot evaluate in terms of objectives which we, as individual extension workers or as part of an extension team, made or will make no attempt to reach. To take an extreme case, how satisfied would you be with evaluation in terms of objectives if not one activity that you carried out could possibly have led the people in the direction you had wished them to go?

We should also realize that analyzing the possibility of reaching an objective through a study of the teaching procedures and subject matter will provide a partial answer to the ultimate effectiveness of our work. By carrying out this preliminary step in evaluation, either when planning a teaching procedure or when planning an evaluation project, we determine whether the objectives were realistic or idealized, material or abstract.

If we find that no attempt was made to reach the objectives, we can anticipate that the work will be ineffective in terms of those objectives. If we conclude that inadequate methods or insufficient subject matter were used, we can expect the effectiveness will be impaired proportionately. If we feel that objectives were properly stated, appropriate actions were outlined, and adequate subject matter was provided, we can anticipate gratifying results in the way of desired behavior changes.

Learning Experiences

In educational extension work, the next step after clarifying objectives is planning for and defining the **learning experiences** to be provided for the learners. Essentially, learning takes place through the experiences the learner has and the reactions he makes to his environment. It does not take place through what you do as a teacher, but through what the student goes through as a learner. There may be only one teacher in the teaching situation but there are as many learning experiences as there are learners participating.

In the evaluation process, we assume that planning for the learning experiences has been carried out, and proceed to analyze the plan. In order to make both planning and analysis more meaningful, we have some principles for selecting and judging teaching activities (program execution). If planning for the learning experiences does not take into account the following principles, its effectiveness might well be questioned without further examination.

1. The learning experience should be appropriate to the level of education, interests, and facilities of the learners.
2. The learning experience should be the considered choice from among the possible ways of accomplishing the desired behavior change.
3. The learning experience should be set up so that the learners have a chance to practice the desired kind of behavior.

4. The learning experience should be such that the learners will get satisfaction from carrying out the desired kind of behavior.
5. The learning experience should lead to interest in accomplishing other desirable behavior changes.

Let us take an example from extension work. An objective in your 4-H Club program may be to help boys and girls in 4-H Clubs to acquire the ability to perform the duties of club officers. Such an objective has such long standing approval in extension circles that it requires no evaluation in its own right. A more appropriate evaluation for your own purposes might be to determine to what degree the club officers are functioning as they should.

You believe that the club president should know parliamentary rules and procedures, and be able to enforce them diplomatically. You believe that he should be poised and confident, and that he should speak clearly and audibly. You believe that the secretary should keep complete and accurate records of meetings and be able to read them clearly and audibly. You have similar criteria for the other officers.

Before you can begin looking for evidence that the officers are functioning properly, you need to do some desk work. Put down and study the activities **you,** as an extension worker or as part of the extension team, carried out. This should include the learning experiences you provided for training local leaders to train the 4-H Club officers.

Did you arrange for the 4-H members to be taught how to do each of the things that you are expecting of them? Did you see to it that the information needed to do these things reached every person who needed it? Was it in terms that they could understand and use? What knowledge, skills, or attitudes ought to change in these boys and girls as a result of the teaching activity that took place? By such thought patterns, we move from "What behavior change does the objective imply?" to "What behavior changes could the experience bring about?"

How did you do your work? **What** did you try to teach? **Whom** did you try to teach? **What** did you expect the learners to know (to be able to do) as a result? Questions such as these fit into the pattern of the preferred way of stating educational objectives so that they include the three components outlined in Chapter V.

Analysis of the Teaching Plan

After you have clarified exactly what your objectives are, what goal you hope the learners will reach, what they can be expected to learn or acquire as a result of the teaching or training, and which ones can acquire the knowledge, you have the background for an evaluation project on your teaching plan. After these are made clear, the other steps in the evaluation process are much easier.

Again we want to stress the fact that evaluation includes not only a measurement of the outcome of your work, but also an analysis of activities contributing to the outcome. This analysis ought to take place during program planning, or during execution, so that urgently needed modifications can still be affected. If it was not done at either of these stages, it must be done as part of the terminal evaluation to eliminate the possibility of undesirable actions being repeated. Because of the importance of evaluation as a learning experience for extension workers, it is essential that only those directly involved in the work do the evaluation.

We have found that a systematic method of carrying out this analysis, following a definite outline, results in the most satisfaction to the evaluator. The following form has worked out very well. The components suggested by the headings are essential; the exact form can be adapted to the wishes and needs of the evaluator.

Teaching-level objectives	Subject-matter	Learning experiences (teaching methods)	People to be reached	Behavior changes expected

The order of these headings follows the sequence appropriate to planning a teaching program, or for analyzing one already set up. First is the matter of developing and stating the teaching-level objectives. The teaching-level objectives referred to here are the extension agent's objectives for the given program or unit, as described in the previous chapter.

Second, for each such objective, certain subject-matter resources are required. List each kind of appropriate subject matter that was used, or that you plan to use, after each objective.

Third, you list those learning experiences which you feel will enable the learner to acquire the subject-matter understanding he will need to make the desired behavior change. Although it may be possible for you to list many different learning experiences after each type of subject matter, list only those that you used, or plan to use. However, you can keep the others in mind as alternatives in the event your evaluation indicates that your choice was not the best one possible under existing circumstances, and there is time to make adjustments.

Fourth, you list the kind of people who can be expected to receive the subject-matter information through each learning experience listed. For example, "those who attend the meeting" are the ones to be listed if some kind of a meeting is the learning experience. "All farmers in (name of) area with an interest in dairy" might be listed if the learning experience is reading a news article or listening to a radio broadcast devoted to a topic like the "weigh-a-day-a-month" plan for improving milk production.

Fifth, the expected change of behavior is listed. At this stage, the most important evaluation of the learning experience takes place. The kind of learning experience limits the kind of changes of behavior that can be expected to take place. Some learning experiences can be expected to change attitudes; others can be expected to change skills. Some can bring about both kinds of behavior change; others may have difficulty accomplishing one.

For example, it would take a homemaker with a very broad knowledge of sewing to learn the skill of making slip covers by listening to a speech on the subject. More than likely, she would be the sort of person who already possessed the skill but just needed to have her memory refreshed on a few points. Most of us would need the added experience of watching a demonstration, backed up by an opportunity to make one under supervision.

When the above analysis has been completed, you have a basis upon which to decide which learning objectives you are justified in studying, and which changes in behavior you can expect to bring about. This systematic method of analysis can be applied with equally beneficial results to a farm visit, a telephone call, or a club meeting. It can be used as an intermediate evaluation of anything from a routine project to your total extension program.

Relationship Between Judging Learning Experiences and Educational Evaluation

Judging learning experiences cannot be carried out without also judging the subject matter used and the kinds of people you expect to reach, as well as relating the learning experiences to the objectives in mind. When this broad conception of judging learning experiences is employed, it becomes a large part of the total evaluation procedure. The part that is left to do is to determine if the expected behavior changes occur, and how satisfactorily.

Stated briefly, judging learning experiences helps you decide which behavior changes to expect; completing the evaluation process helps you determine which expected behavior changes took place, and how well.

Summary

This and the succeeding chapter are devoted to intermediate types of evaluation of special use in extension teaching. In this chapter we have discussed how evaluation enters into judging the probable effectiveness of a plan of work before there are any results to use in the assessment.

In evaluation of this variety, we are concerned with assuring a realistic approach to the problem at hand by stating objectives in a meaningful manner, setting up the action to be carried out, and designing suitable learning experiences. Evaluation is in terms of determining how appropriate the methodology is for accomplishing the objectives outlined.

We have presented a five-part outline for evaluating an extension plan of work. It includes appraisal of the teaching-level objectives, the subject matter to be used, the learning experiences to be set up, the people to be reached, and the behavior changes to be looked for. At this stage we have to compare the elements of the plan as outlined with approved extension practices. Later, when the plan has been executed, we can evaluate the program in terms of observed behavior changes brought about by the action.

References

1. Herzog, Elizabeth. "How Much Are They Helped?" **Children,** November-December 1958. Government Printing Office, Washington, D. C.
2. Tyler, Ralph W. **Basic Principles of Curriculum and Instruction.** Syllabus for Education 360. Chicago, The University of Chicago Press, 1950, Chapter V.

Chapter VII

EVIDENCE OF PROGRESS TOWARD OBJECTIVES

Mrs. Laurel K. Sabrosky

You are well on your way with your evaluation plan after you have selected and defined the objective for evaluation and the work you did to reach it. These lead you directly into determining the purpose of the specific evaluation you wish to carry out, with the necessary clarification and limiting of it.

The next step is to determine what **evidence** you will need to look for so that you will (1) know the situation, and (2) know that you are going in the right direction, or (3) know that your objective has been reached. This is true regardless of the evaluation procedure you will follow (use a record form, make a home visit, carry out a formal survey, or others).

Evidence

Our use of the word "evidence" is in accord with dictionary definitions: An outward sign, an indication, that which furnishes any mode of proof.

How can you tell that you have made progress, or that you have attained the goal toward which your objective was aimed? Education is successful only to the degree that it has caused changes in the "right" direction. What is the evidence (the outward sign, indication, proof) of "right" status change or progress?

Evidence is not always easy to come by. Human beings are complicated creatures and changes occurring among them are not always obvious to the unaided eye. This forces us to look for evidence of the success of our objectives in terms of observable behavior, or results of behavior. (e.g. Learning how to make a dress: Change in behavior; the dress itself: Result of behavior.)

Behavior, when considered as a **result** of education, is often not in terms of readily observable physical action. However, we can find evidence of it by looking for improvement in skills, better understanding of concepts, increased ability to solve problems, changed attitudes, appreciation of different things, shifts in values, changed interests, increased knowledge, adoption of approved practices.

To select the types of behavior to measure, you must first check the types of change you have been or will be trying to help bring about. This involves reviewing your plan of work. Different teaching methods or techniques have different types of behavior change as their objectives. The important thing is to ask yourself constantly, "Was this change supposed to occur because this particular objective is in the teaching plan? Were these methods the kind that could bring about successful accomplishment of this objective?"

A radio talk may be aimed at producing a desired action or at increasing the knowledge of the listeners. A recipe in a newspaper may be put in for the purpose of encouraging people to use food more economically, tastefully, or nutritionally. A radio talk, news article, or project lesson may be aimed at teaching why something is important. One method may attempt to bring about a change in practice; another may attempt to bring about a change in knowledge, attitude, or appreciation. We shall need to decide which and how many types of behavior we want to measure.

Depth of Evaluation

1. **In Terms of Changes in Behavior of People**—One level at which to measure your progress or attainment as a helper, educator, teacher or organizer is to find out which changes have taken place.

   a. **In the People Themselves**—Have they (the farmers, homemakers, young people) changed their attitudes or skills; have they done anything as a result of the extension activity or method; has their status changed in any way since the extension action was started?

   b. **In the Extension Workers** (the county agents, local leaders, or officers) being taught or trained so that they can help others learn something new.

2. **In Terms of Opportunity**—When it is difficult or impossible to measure progress at the level of original status or change in people themselves, it is desirable to measure work in terms of the **learning situation** we have set up. If no learning situations are set up (no written materials go out, no talks are given, no demonstrations are put on, no visits

are made), we cannot expect the people to learn anything as a result of extension work.

It is logical to assume that meetings which have 1,000 people in attendance have a greater chance of changing behavior than no meetings at all. However, it would depend on the learning situation whether a meeting with 1,000 in attendance, 10 meetings with 100 in attendance at each, or 100 meetings with 10 in attendance at each would provide the best learning opportunity.

In general, the more opportunities you provide, and the better they are, the better is your chance of accomplishing your extension objective. A well organized and well attended 4-H Club event should offer a better opportunity for boys and girls to learn than a poorly organized and poorly attended one.

By appraising the soundness of organization, and the attendance, coverage and subject matter, you have some information about possible success. You still have no real qualitative measure of educational change in behavior of the people who are reached, but you have a good idea of quantity.

Levels of Behavior Change at Which to Measure

There are three convenient levels for doing evaluative measurements. The first one is before any change, or any further change, occurs (the existing behavior position before any extension teaching has been done, or before more teaching is undertaken). This point, level, phase where the learners are before the plan for changing behavior goes into action is referred to as the **benchmark.**

The decision whether or not to establish a benchmark has to be made at this step in the evaluation outline. If you want to know what progress has been made, or will be made, you should have some idea of the situation at the benchmark. If you know pretty well what it is, there is little need to establish it formally. It is not necessary if you are interested mainly in whether the goal is reached, and not in how far learners had to progress to reach the goal. However, by establishing the benchmark, you provide yourself with a good starting point for program planning by finding out how urgent the need for change really is.

The second level is during the progress of the learners toward the ultimate goal of the objective. Sometimes we think of evaluation at this level as a **progress report.** In evaluating a long-time objective, progress reports are necessary to tell us how things are getting along. They serve to tell us if the teaching methods are effective, which ones are most effective and which ones need to be emphasized in the program of work if the objective is to be reached. They tell us which steps have been accomplished and which remain to be accomplished. Also, they may suggest other steps which should be inserted because of unexpected developments during the action.

The third or **attainment-of-goal** level is at the close when we must determine whether or not the final goal of the objective was reached. When you find out if few, some, or all of the people reached the goal, you know whether to retain the objective in the plan of work, redesign it and try again, or substitute another one. You can find out whether teaching methods have been effective or not, and under what conditions.

Although it is desirable for most situations to check at each level, two of them may be incorporated into a single testing program. Because it is necessary in most plans to establish a base for measurement, the evaluation plan may call for measurement at two levels, the benchmark level and either the progress level or the attainment-of-goal level.

Factors for Deciding Which Evidences of Behavior to Use

Our next decision concerns which changes in people are to be considered as evidences of status, progress, or attainment. There are several factors which need to be considered in this regard. Most of them revolve around problems of measurement or ways to simplify measurement.

1. **Which Evidences Will the Learners Have an Opportunity to Display?**—Whether or not the learners have the opportunity to develop a new kind of behavior, and to display it so that it can be detected, depends on two conditions. The first one is that the learning experiences must have occurred. We discussed the relationship of the learning experience to behavior change in Chapter VI. The learners must have had a chance to learn something before it is reasonable to look for evidence that they did learn it.

All of the people who had the opportunity, or those you wish to have the opportunity in the future, to learn enough to make a change in behavior, become the **population** from which you collect evidence. How to sample this total population, to make evaluation procedures manageable, is discussed in Chapter X.

The second consideration is whether or not the people had a chance to display the change of behavior. This means that it is necessary to decide if the people have the facilities, the need, and the motivation to make the change of behavior a part of their thinking or living habits. Only where a favorable or permissive climate for such changes exists is it meaningful to look for evidences of teaching success.

This factor limits the evidences much more in connection with youth work than it does with adult education processes. It is true that favorable learning experiences can be set up with equal facility for youth educational groups (the schools, YMW work, 4-H Clubs) or adult education work. However, a favorable climate for displaying evidence of behavior change is often long delayed in connection with youth education work.

2. **Will the Things You Can Count Be Adequate as Evidence?**—These outward signs are the easiest results to use in evaluation, and they make much better working materials for measurement techniques. However, if you do not consider these types of evidence as adequate proof of status, progress, or goal attainment, you will have to think about ways of assessing intangible evidences, such as changes in knowledge, attitude, or thinking.

3. **What Is Your Own Facility in Measuring Different Types of Behavior?**—There is no question about the fact that it is easier to measure changes of behavior which have tangible evidences for proof than it is to measure those which have intangible evidences. The fact that a homemaker has made use of a recipe she has seen in a paper is easier to determine than the fact that she now has a better understanding of the value of good nutrition in child development. The fact that a family gathers around the kitchen table to plan the duties of each member is not difficult to ascertain. But, whether or not the family now gets along better as a result of this planning experience, is an intangible result that is hard to measure.

To determine the real results of extension teaching, we cannot ignore intangible results just because they are hard to measure. On the other hand, while we are becoming accustomed to evaluation procedures, it is highly desirable that we become familiar with the techniques, methods, and procedures by measuring those results which are easier to measure. When the techniques are mastered, we can use them to measure those types of results which are more difficult to measure.

Just as we learn to read by starting with simple words, simple sentences, and simple stories, so we learn evaluation methods by starting with the simple and advancing to the more difficult measuring devices. Keep in mind, however, when you begin with the easier methods, that you do not get the idea that all evaluation is concerned only with tangible results that can be seen and counted.

4. **How Will You Limit the Number of Evidences to Look For?**—If the list of evidences is very brief, and to look for all of them would not require a great deal of work, all of them should be included in your plan. If the objective to be studied has been defined and limited so that it requires only one or a very few actions or changes in behavior to reach it, those are the ones, and the only ones you need to look for.

More often, however, there are numerous evidences of change of behavior. When such is the case, there are three frequently used ways of selecting which ones to use in your evaluation.

a. Study the list of evidences to see if there are any readily apparent types of changes that, by occurring, would serve as clues to the fact that other changes must have been brought about. Sometimes only these **discriminating evidences,** the ones which can be taken as indicators of others, are enough to look for. In a progress report, however, be careful that those actions which are selected as indicators of others are adequate to establish status of steps in progress.

b. If there are no such discriminating evidences, or if there are too many of them, study all observable changes to see which can serve as the most **important indicators** of change in behavior. If you cannot look for all changes, certainly those you consider most important will be of most value to you.

c. If your long list cannot be reduced by either of these methods, you can reduce it arbitrarily by **random selection**—choosing at random a practical number of them for careful study. Random methodology will be discussed in Chapter X.

5. **What Will You Do About the Time Factor?**—Seldom is a change of behavior, whatever the type, apparent immediately upon reception of the teaching stimulus. For certain types of change, such as in attitudes, appreciation, knowledge, or values, a change in thought processes may take place immediately. However, the learner will need time to think over the material, weigh it in his mind, and be exposed again to his customary sources of guidance, before a change of any permanence takes place.

Knowledge which is evidenced immediately after presentation of new information may be forgotten just as quickly. When the desired change in behavior is a change of practice, the time element is extremely important. Some things cannot be done until, for example, a certain season has arrived, certain materials are assembled, money becomes available to buy what is needed, or the family is all together. Some things, such as remodeling a house, take a long time to do. Some things take place in stages—planting the garden, harvesting it, canning the produce, serving balanced meals throughout the year.

In planning the evaluation, therefore, it must be decided how long it is reasonable for you to delay your decision as to whether or not a certain practice has been adopted. Judgment must be used in this phase. We do not know exactly how long we should wait after a person has had a chance to carry out a practice, and does not do so, before deciding that he is not going to do it. The longer he refrains, certainly, the less likely it is that he will ever carry it out.

On the other hand, there is a limit to the time we can wait and still expect to find out exactly what was accomplished. For daily happenings, 2 or 3 days may bracket the memory span. Regarding seasonal activities, we might be justified in expecting a respondent to remember an important happening during the rest of the year.

When reporting memory items, it should be made clear that the data are from memory and not from records. The passage of time distorts accuracy of recall considerably in excess of a straight line relationship. Periodic observations spread over the period, rather than a single terminal observation, will improve accuracy of recall. For some types of information, you may find it best to collect data at different times, because the various actions or desired behaviors may be affected differently by time.

6. **What Decisions Must You Make Regarding Units?**—Will the evidence be by degrees of attainment, amounts of things accomplished, or length of time an action has been carried out? Will evidence of a year-around garden be in terms of certain vegetables raised, amounts of certain vegetables canned and stored, or in other terms? Will it be the average number of quarts of vegetables per person, or will it be the percentage of the homemakers who can a recommended average or minimum amount of vegetables?

What will be the evidence you want or need to form a basis for your analysis? What will be the units? Will they be comparable, or will you have to convert them before you can make comparisons?

7. **What Face Data Will You Need?**—We think of those items of descriptive or background information needed from or about all individuals from whom we get our evidence as **face data.** They include such information as years of schooling, age, sex, size of farm, number of children in the family, income, and so on. Face data are used for three major purposes.

a. **Classification of Respondents**—Usually it is desirable to find out if there is a difference in the replies or reactions of individuals possessing different degrees of certain characteristics. This requires an analysis of the replies using the classification system set up for a particular characteristic as a sorting device. For example, if you want to know if there is a difference in replies among those who have had different amounts of formal schooling, you would need to include a face data question on number of years of formal schooling. If part-time farmers are to be compared with full-time farmers, appropriate face data questions must be included to provide that information.

b. **Purposes of Standardization**—To compare the findings from one study with those from another, the individuals in both studies must be comparable, at least in terms of the sort factors under consideration. The findings in a study of commercial farmers are not directly comparable with findings in a study of subsistence farmers. If either or both studies have no face data questions on size or type of farm, the findings from the two cannot be standardized for purposes of studying the relationship of size of farm operation to other aspects under consideration.

c. **Purposes of Comparison**—Frequently it is desirable to compare data from a study of only a part of the population with data from a study of the whole population (universe). Usually census data are used for such purposes because they deal with the whole population and because they are about the best source of depth or trend information. In such cases, face data questions comparable to those used to secure census data should be included.

For example, if the study data are to be checked with census data on age, residence, or similar classifications, face data questions must be included which will provide data suitable for assembling into the categories used by the census. When comparisons are made between studies or with census data, it is vital that data be secured in sufficient detail so that they can be converted to comparable units or reporting intervals, although they do not have to be collected in identical form.

This means that in planning an evaluation, and especially while constructing questions, the analysis which will eventually be made of the results must be given serious thought. Think over the teaching methods and the informational materials. What do you want to know about the behavior changes you are trying to bring about? What human reactions do you want to explore?

Do you expect certain types of people to react more favorably than others to your teaching? Do you expect that certain mental or material possessions or qualifications will make a difference? If you do, be certain to include appropriate face data so that you can sort on these characteristics, and evaluate their effects properly.

Face data will help you ascertain differences, make comparisons with other studies, and check the probable accuracy of your data. All of this contributes to your major purpose in evaluation, truing your work and making new and better plans for your work.

Summary

In this chapter we have probed more intensively into the problems of evaluating during the progress of a program. During this dynamic stage we must look for evidences or results which we can use to assess the progress of our work. We might prefer to look for them in terms of changes in people, but, because of the delay before most changes actually occur, we might have to settle for counting opportunities provided for changes to occur.

We suggested three stages when evaluation is in order, referring to them as the benchmark, the progress report, and the attainment-of-goal levels, and suggested some techniques appropriate to each. We then devoted the balance of the chapter to discussing seven factors to consider in deciding what evidences to use in evaluation. We gave special attention to the last factor, the matter of face data, because of its importance to the more detailed types of evaluative action we will be discussing in forthcoming chapters in this section.

References

1. Smith, E. R. and Tyler, R. W. **Appraising and Recording Student Progress.** New York, Harper and Brothers, 1942. Chapter V.
2. Tyler, R. W. **Basic Principles of Curriculum and Instruction.** Syllabus for Education 360. Chicago, The University of Chicago Press, 1950. Chapter II, pp. 15-34.

Chapter VIII

IDENTIFYING PROBLEM AREAS AND STUDY OBJECTIVES

Ward F. Porter

In this chapter we return from our exploration of the various types of evaluation to a brief accounting of the initial phase of the evaluation process. This chapter is aimed at directing our thinking to the "big picture" of evaluation planning, before moving on to specific techniques of evaluation.

Our immediate objective is to help the reader develop a better understanding of the process whereby appropriate research or evaluation problems are identified and study objectives determined. We hope to emphasize the urgency of doing a thorough job of understanding the problem before starting to evaluate it. All too frequently, the importance of this first step is not properly appreciated nor sufficiently understood. The net result it that much needlessly slipshod evaluation is turned out and, unfortunately, gets used without question in policy making and program planning actions.

The Importance of Problem Identification

Precise identification of the problem needing investigation is essential from many standpoints. This step helps the extension worker clarify his thinking and facilitates the selection of suitable study objectives. It also serves as a basis for sound decision-making regarding evidence to look for, data to be collected, and analyses to be made. It sharpens up his interpretations and suggests ways of getting the results into use.

The need for evaluation and research in Extension grows out of our role as educators and the resulting necessity for determining the effectiveness of our programs and methods. Only in this way can we make necessary and realistic modifications in future actions to assure maximum benefits from our extension teaching. In addition, all extension workers are confronted with the problem of identifying the needs and interests of their clientele.

Identifying these needs becomes an objective of sufficient importance to demand the most careful evaluative and research techniques. Although some situations are such that relatively informal exploratory devices will serve, others may require full-blown formal studies. Whatever measures are used, the results will be applied in determining action programs, in deciding upon appropriate methods, or in establishing benchmarks by which educational progress can, at a later date, be measured.

Selecting a Problem for Study

The fact that you have identified a problem area does not signify that you are ready to begin your evaluation. More often than not, the problem you have selected is much too all-inclusive to be handled adequately in a single evaluation. There is still the matter of identifying the **specific** problem for your evaluation.

There are many factors that need to be considered before selecting our specific problem area and presenting it in a form suitable for evaluation. We must consider our own personal and professional needs and interests in the matter, as well as the needs and interests of our associates, our lay leaders, and our clientele. Obviously, there will be some overlapping of interests and concerns, as well as occasional conflicts. Any differences, of course, must in some way be resolved.

We must decide not only the **general** area of concern, but also the **particular** phase or aspect that is to be studied. As an example, let us take the case of the extension worker who would like to know how effective he has been in reaching a certain group of families with a particular program. Already he has to a degree, limited his area for study by specifying a broad problem area. Before he has a study he can cope with, however, he will have to narrow it down further. He might choose to select acceptance of a single practice as an indicator of how well he has succeeded with his educational program. In short, if his program involved pasture improvement, he might select just one of the several recommended practices—periodic clipping, for example—for the detailed study.

The decision as to which problem is to be investigated at any one time inevitably must require some judgment (evaluation) as to the relative importance of the many different problems. The agent might try outlining worthwhile areas of study, some problems in each, and the specific topics which could be studied under them. He might organize them into some form of priority listing in terms of their probable importance—to himself, his associates, his lay leaders or his clientele. Then, by adopting the policy of "first things first" as in program planning—he can decide on his plan of action.

Such a plan should be executed with care, however, as failure to correctly assess the relative importance of the many problems demanding investigation can result in frustration and discouragement. This is another way of saying that evaluation, like most things that are worthwhile, requires considerable planning.

Time is another factor which must be given adequate attention. An all-too-frequent criticism of evaluation plans, and research in general, is the inadequate amount of time allotted for the kind of investigation that is outlined. Adequate attention must also be given to other types of resources which must be employed in the study, such as personnel, finances, skills and know-how, equipment and facilities, and related items.

Unfortunately, these resource factors are not always considered carefully enough in planning an evaluation or survey. Often the result of such oversight is a study that never gets finished, sometimes because funds are expended before the data can be analyzed. Another distressingly frequent occurrence is when researchers, with an impressive array of electronic equipment at their disposal, dash into the field with a carelessly constructed, "shotgun" type questionnaire which is utterly incapable of securing valid, reliable, and objective data.

One other important preliminary consideration is the amount and kind of information already at hand that may be relevant to the problem. Solutions, or at least objective information from which inferences can be drawn, may already be available. In this event, duplication would be a needless waste of time and other resources.

One note of caution is in order in this connection. Inferences drawn from other studies or research findings of presumed relevance should not be applied to the new setting without thorough checks for comparability. Overgeneralization of research findings, without regard for comparability of situations, people involved, research techniques used, and related factors, is altogether too common, and sometimes fraught with serious consequences.

Selecting and Defining Study Objectives

Once the problem area has been identified, it becomes a matter of pinpointing the specific objectives of the forthcoming study. In a sense, the identification of the problem(s) may be thought of as establishing the overall **purpose** of the investigation. That is, a study may be undertaken to establish benchmarks for later program or method evaluation, to indicate needs and interests of the clientele, or to establish an estimate of the effectiveness of a meeting in changing human behavior. On the other hand, the proper statement of study objectives indicates specifically what aspects of the problem are to be emphasized.

We can illustrate this distinction between identifying the problem and setting up the objectives for study purposes by this example. A home demonstration agent might have the problem of evaluating a leader-training meeting. Her study objectives might be as comprehensive as the educational program involved at the meeting. At the same time, she might quite properly limit her evaluation, because of other pressing demands on her time and other resources, to a single aspect of the program. She might, for example, be content with determining the number of leaders attending the meeting who learned how to freeze peaches, even though freezing procedures for several other foods were discussed.

It is important in this connection to distinguish between educational objectives and study objectives. Previously, we defined educational objectives in terms of changes in human behavior (e.g., an increase in knowledge, understanding, or skills, or changes in attitudes or opinions). Study objectives may be phrased in terms of educational objectives, but do not have to be, depending upon the nature of the problem.

In a strictly situational type of analysis, such as a countrywide survey, no teaching objectives as such need be directly involved. In an evaluation of the effectiveness of an educational effort, however, study objectives most certainly would have to be consistent with the program's specific educational objectives. In either case, of course, the scope of the investigation may be limited.

The study objectives also can be stated in terms of learning opportunities provided. Instead of attempting to appraise knowledge or skills acquired, the evaluation can be limited to determining the number of exposures to the learning situation—how many attended the meeting, how many heard the radio program, how many made the field trip, and so on. As indicated earlier, such opportunity level evaluations have their shortcomings and limitations.

Stating Study Objectives

The factors mentioned previously as affecting problem identification, such as resources available, are all equally relevant in selecting, defining, and amplifying study objectives. Here again, it may be helpful to list the various potential study objectives, as encompassed by the problem area selected, in the order of their importance and then consider each in the light of resources available.

We can help ourselves a lot when it comes to collecting, analyzing and interpreting our data if we do a good job of stating our objectives. The more precisely we state them the easier it will be to design questions that will secure the information we need.

One device for gaining precision in our statement of study objectives is to develop a series of basic questions around which the investigation can be focused. In a study of the effectiveness of a marketing-information-for-consumers (MIC) program, for example, one of our study objectives might be: To determine the nature and scope of the MIC radio audience in "X" County.

We could bring this objective into sharper focus by listing several basic questions that we want answered. Such a list might include questions of this order: (1) How many farm families in "X" County have regular access to a radio that is in working order? (2) How many of these families (one or more members of) listen regularly (over half of the broadcasts) to the MIC radio program? (3) What proportion of these families live on farms? In the open country? In small towns or villages? In cities (over 10,000 inhabitants)?

In the more formal and comprehensive types of research projects, researchers follow up their statement of objectives by working up appropriate hypotheses relating to them, which they proceed to test by statistical and other methods. These hypotheses are their considered opinions, based on inferences of what they expect their research will find out regarding their objectives. The hypotheses are then tested by the data collected and either supported or rejected, in part or in whole.

The hypothesis becomes the focus of the investigation, directing the attention of the researcher toward certain types of facts. By so limiting or delineating the study, the random and unsystematic search for facts is avoided.

An example of one hypothesis that might be tested in an extension study is the following: Result demonstrations are more effective than the mass media in promoting the adoption of certain recommended practices among low-income farm families of "X" County. To confirm or deny this hypothesis, we would have to work up a series of questions to help us determine, (1) the degree of effectiveness of the result demonstration, and (2) the effectiveness of the mass media, both separately and collectively. First, however, we would have to select suitable practices and devise some criterion of effectiveness—an exposure index, an acceptance score, or similar device.

Only when the study objectives are precisely stated are we in a position to plan and carry out the other phases of our evaluation or study in an effective manner. Time and effort devoted to this initial phase of the process will prove very rewarding in terms of the quality, reliability and usefulness of our results.

Summary

In this final chapter before elaborating on the techniques for assembling and analyzing evaluative data, we have briefly discussed ways and means of selecting and stating our evaluation problem and study objectives. In essence this is largely a matter of identifying problem areas of major interest and concern, selecting one or more for immediate investigation, determining our study objectives, and pinpointing the major questions or hypotheses that will serve as the focus of the study. Although simple-sounding in outline, each step merits careful and detailed consideration. The success of the entire evaluation process will depend on how well each step is worked out.

References

1. Good, C. V., and Scates, D. E. **Methods of Research,** New York, Appleton-Century-Crofts, Inc., 1954.
2. Parten, M. B. **Surveys, Pools, and Samples,** New York, Harper & Bros., 1950.

Chapter IX

SOURCES OF DATA

Darcie Byrn

As we stated in Chapter VII, it is very common for the evaluation process to start off with a benchmark phase. The need to know where you are so you can determine later how far you got is fundamental to evaluation theory. To set up this benchmark, we must learn what we can about the existing situation.

If we are old timers on the scene, we can use our own experiences as a source for much of our setup data. It might also be possible to call on others to supply much of the information we need on the basis of their first hand knowledge of the situation. However, even the most observant person needs information of a factual or statistical nature to supplement his first hand acquaintance with the setting.

For most of our purposes we will find, upon investigation, that there is a great amount of material already assembled which will be of use to us, for purposes of control and comparison if not for actually solving our problem. By using these sources wherever and whenever they can help with our evaluation we are saving valuable time and other resources for use in later stages of our evaluation or for further evaluative actions.

Primary or Secondary Source

Traditionally, sources of data have been classified into two categories, primary and secondary. Our major interest in this chapter is with the latter kind; assembly of data from primary sources will be the central theme of the remaining chapters of this section.

We can think of **primary** sources as the original documents, the first reporting of the facts, the first grouping of the raw data. **Secondary** sources bring together facts from primary sources. They can make a substantial contribution to knowledge by way of analysis and interpretation of primary sources, and by making information presented in obscure primary sources more readily available and understandable, but they do not make the original contact with the unknown.

Data assembled by means of questionnaires, schedules, interviews and similar instruments are data from primary sources. Records from 4-H members, result demonstrators, farm and home development cooperators, and periodic reports of extension personnel are primary sources of data. Summary reports of State and Federal extension officials occupy a somewhat uncertain position in this classification. In one sense they are original tabulations and compilations; in another they are compilations and interpretations of data from original sources.

Often we will find it difficult to make the decision as to whether a source can be called a primary or a secondary one. For instance, reports compiled from census interview forms are primary sources. Reports compiled from census reports, irrespective of whether they were done by census personnel or people in other organizations and agencies, must be classified as secondary sources. Many of the reports we in Extension make or help to assemble contain both original and secondary data. We will soon discover, however, that we can make good use of either kind by making proper adjustments and allowances.

Although this decision is purely academic in most situations, it must be taken into consideration when evaluation is being planned. It will guide us in deciding how much trust to place in the data through our ideas of the presumed professional stature of the author. For instance, most of us accept original research prepared by State and national governmental agencies and established research organizations with respect, if not without question or qualification. At the same time, we would want to make a personal decision regarding the qualifications of a committee of club members or a masters-level graduate student before accepting their findings without question.

By the same token, we would want our secondary source to be at least as qualified as the primary source it quotes. Much good primary research gets abused during the process of inefficient secondary analysis. Similarly, more than one presumably sound research organization has committed the error of attempting to salvage inadequate original research by applying elaborate and unjustified statistical manipulations while synthesizing results of the original study.

There are reasons for preferring a primary source over a secondary one when both are available. In the case of the former, there are fewer opportunities for errors to get into the final manuscript. Counts, sorts, tabulations and interpretations are made by those most likely to understand the material. When special definitions are necessary, there is greater likelihood that they will be used consistently throughout the manuscript. Groupings will be made in terms of the original material, rather than the needs of the secondary researcher who may never have an opportunity to examine the original data.

Secondary sources may contain mistakes of both transcription and interpretation resulting from the extra handling and lack of familiarity with the original data and the methods by which they were assembled. Quite often there is a tendency to generalize or extrapolate (go beyond the data) without adequate justification. This is not to suggest that secondary sources of data have no value, merely that they may have limitations less likely to be encountered in the original source.

Secondary sources may point out important primary sources which we might never learn of if left to our own resources. They are helpful in opening up an area and in suggesting directions for individual exploration. They show us where we can locate original sources and what we are likely to find. Their interpretations may help us to understand the data of the original report. If properly used, they can save us a great amount of time and effort without leading us astray from our objectives.

Sometimes data from either source are sufficiently comparable so that they can be used to answer our questions, and further research on our part will not be necessary. If either is used in place of personal research we must be certain to check on reliability, accuracy and applicability as thoroughly as we can. This is an occasion when we can draw on our own family of experts—the research and subject-matter people at the land-grant institutions and in the Federal Extension Service offices—to help us decide on quality of sources.

Useful Data Sources for the Extension Worker

No extension worker need fear that he will have to "go it alone" on a research problem. More than likely, he will be overwhelmed by the wealth of background material that can be made available to him once he gets tuned in on the proper communication channels.

It is difficult to conceive of a problem that will lie entirely outside the area of existing research. Seldom will you find a piece of research which exactly fits your problem, but you can be assured that you will find several close enough to give you a lot of help in laying out your own study. If you need help with pulling these leads together, there are persons qualified to advise you in the State Extension office, possibly even in the county, and surely in the Federal office.

Sources of both published and unpublished material are too extensive to be given comprehensive treatment here. However, we will suggest some favorite starting points for the exploratory phase of an evaluation assignment.

General—In our work we are likely to think first of the U. S. Department of Agriculture. The vast research resources of the Department are available to us at little or no cost to us personally. Our contact for published materials of the Department is its Office of Information or the information outlet for the appropriate agency within the Department.

We should become familiar with its **Bibliography of Agriculture,** published monthly by the departmental library. This contains listings of many of the current agricultural publications, both foreign and domestic. Incidentally, this is a good institution to become acquainted with if you are in a position to come to Washington, or can make use of its loan facilities.

Another good starting point is the monthly publication of the U. S. Government Printing Office, **Monthly List of Government Publications.** It has extensive listings of agricultural items and indicates where free copies might be available, or else quotes prices if they must be purchased. In addition, several of the agencies have periodic listings of materials prepared within the agency. Often it is possible to be placed on their mailing lists to receive such notices regularly. Very often, free copies are available for distribution, at least for a short while after publication, from their information branch or division.

The Department or its agencies assembles a great many periodic and special purpose reports providing a vast amount of trend, estimating, production, predictive and regulative information on agricultural subjects. It maintains research experts, consultants and technicians in the Washington offices and dispursed throughout the States and Territories, and has many others on foreign assignment.

The resources of the U. S. Department of Commerce are also available to us. We can make a great deal of use of the reports of its Bureau of the Census. To learn what this agency has to offer, we should know of its **Quarterly List of Census Publications.** It lists virtually everything issued by the bureau and tells you where things are available, together with the price if there is a charge. The U. S. Department of Labor also does research of interest to us. We can get valuable information on the agricultural labor force, the migratory labor force and like information from them.

At the State level, we have the numerous facilities of the land-grant institutions (the Extension Service, the Experiment Station, the subject-matter departments, the agricultural college and experiment station libraries) at our disposal. Their published works generally are available without cost to residents of the State from their Agricultural Mailing Room or some central office similarly titled.

Many of the regional and district people are only too happy to help us. Much service can also be rendered by the State Departments of Agriculture. Then, too, there are the larger farm organizations, many of whom have their research and public information functions well established at both State and national levels.

Locally we have the facilities of the County Extension Office, where many of the published materials originating from the above mentioned sources are already available. We can call on the local civic organizations and local chapters of religious and professional organizations. Then, too, there are the other civic officials around the Court House and elsewhere in the towns and cities of the county, as well as the elementary and secondary educators and education officials of the county.

Home Economics—Most of the general sources already mentioned will give us leads to materials with special application to home economics evaluation tasks. However, we would also mention several agencies, or branches and divisions within agencies of the U. S. Department of Agriculture, whose task is developing knowledge about home economics.

Research in phases of home economics has been carried on by a succession of agencies within the Department since 1894. In the present departmental setup, the work is being carried forward by the Institute of Home Economics in the Agricultural Research Service. Special phases of the work are assigned to the Clothing and Housing, the Household Economics, and the Human Nutrition Research Divisions within the Institute.

Home economists also may be found in other agencies where, although they may carry any one of several social scientist labels, will be working with the problems of the homemaker. Our own division employs home economists in its Federal offices and the agency maintains a separate division, the Home Economics Programs Division, to advise on and conduct home economics research, among other things. Published works of all these persons will be found listed in the sources mentioned previously.

Other places where we should cast an interested glance from time to time, because of the valuable research they do in related areas, are: (1) The U. S. Department of Health, Education and Welfare, where they do so much work related to the home (rural health, child welfare, pure food and drug work, rural education); (2) The U. S. Department of Labor, for its studies of women in industry (and out of the home), and child labor problems; (3) The U. S. Department of Commerce and its Census Bureau with its many Agriculture, Housing and Population reports.

As they do at the national level, most of the organizations mentioned earlier have home economists doing home economics research at the State level. The land-grant institutions have schools or departments of home economics where they train home economists and do home economics research.

Most of the counties have home economists on the county extension staff to head up the home demonstration work and to participate in the farm and home development work. Most of these people report on research they have done from time to time, and the published reports became available, at least locally, for use of their fellow workers and their clientele.

Extension—We have already mentioned some sources within Extension but we would like to suggest a few more. One source, actually several sources, of new information on research are the newsletters of the various extension divisions and branches to their counterparts at other levels within the agency. Most of us at least have access to them, if we do not receive them directly.

One of the special assignments of the Division of Extension Research and Training is passing along knowledge of research findings. To this end they put out a series of periodic reports of use as material or idea sources.

They prepare annual digests of statistics compiled from the annual reports of county and State level extension workers. Each year they publish **Review of Extension Research,** which is a compilation of summary reports of research related to organization, functioning and administration of the work which were completed during the previous year.

At five-year intervals they publish **Bibliography on Extension Research,** which is a listing of extension research reports published or reported on during the previous five-year period. To supplement these two reports they prepare special listings from time to time as conditions warrant. Recent topics for special listings have been radio, television, and communications research.

Recently, the Division has begun preparing a series of popular presentations of results of extension research. This numbered series is called, **ER&T Research Summaries.** Distribution is limited; presently the Federal staff and the State leaders of extension studies are receiving all summaries. However, interested persons can get separate summaries upon request.

Local workers will find good source materials in their office files of annual reports and in their records of farm and home development work and rural development work. Similar use can be made of 4-H record books and result demonstration records.

Summary

If we are beginning researchers (evaluators) we need not be overly concerned about background data for our problem. We just have to know how to get started at digging it out and then direct our energies to being selective of what we find so that we use only that which is proper and relevant to our work.

In this chapter we have suggested that sources of data are of two varieties, primary and secondary, and have stated a qualified preference for the former. We also stated that many research reports contain both kinds of data and that often the classification is difficult to make. We went on to say that, rather than make the distinction for classification reasons only, we should make it for purposes of determining how freely we should use the material in our own problems.

We mentioned some favorite ways of getting started on the task of locating background material for evaluation work and suggested sources at different levels where research results are made known, or where we can direct inquiries about specific topics. We gave some leads for securing benchmark and other data for general agricultural extension problems then suggested some for home economics problems. We closed by mentioning some specific materials which are available from the Division of Extension Research and Training.

Chapter X

SAMPLING

Mrs. Laurel K. Sabrosky

Extension workers usually reach and teach more people than they can hope to get evidence from to determine whether their clientiele has need for, or has experienced, an educational change. In a way this is unfortunate, as the way to get complete and reliable information is to go to all the people and find out from them, in one way or another, whether they need or have learned the things you are trying to teach. Since this is seldom possible in extension work, we have to get information from a **part** of the people about **some** of the things taught.

When you choose only part of the people from whom to get information, you are **sampling.** Since the methods we use in selecting the sample affect so greatly the reliability of the information we assemble, we need to have a fairly detailed understanding of how sampling is done.

Even though very few of you may ever have occasion to develop a comprehensive sampling plan, you do use sampling methods more often than you may realize and regardless of how simple your data collection effort may be.

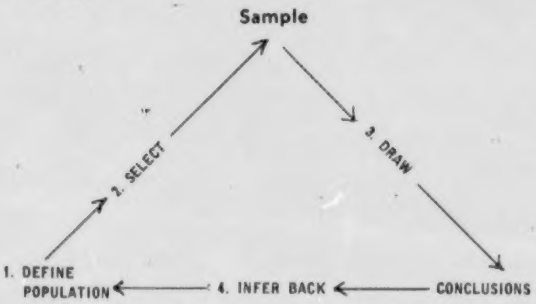
This means that sound principles of sampling are important to all of us in Extension. Proper sampling principles and procedures should be in the mind of every extension worker whenever he is planning or evaluating his work. In the balance of this chapter we will discuss some approved sampling procedures which have found a place in extension work.

General Definitions and Principles

Sampling is selecting a small number of units from among the whole group concerned in such a manner that they can be used to make estimates about the whole group. Too often, a sample is thought of as consisting of only the first part of this definition: " . . . selecting a small number of units from among the whole group concerned . . . ". However, to be meaningful in statistics, and in any evaluative effort, the sample must be drawn in such a manner that the entire definition applies. It is essential that " . . . in such a manner that they can be used to make estimates about the whole group" enters into plans for drawing the sample.

Four steps are involved in sampling, regardless of how elaborate the actual execution may become. They consist of:

1. Defining the population.
2. Choosing and executing the sampling plan.
3. Drawing conclusions based on information extracted from the sample.
4. Inferring conclusions back to the total population.



The accompanying diagram illustrates this sequence of events. The first two steps must be involved in the planning stage; the other two constitute the analysis and interpretation stage. The soundness of the sampling plan will determine how faithfully we can generalize from our sample back to the population.

The **population** for our purpose is the total units to be involved in our evaluation. In extension teaching the units are generally people (homemakers, club members, project participants, potential listeners or viewers, farm and home development families) but in statistical considerations they can be virtually anything (trees, fruit flies, guinea pigs, supermarkets, lumps of coal, fish, cans of corn). Thus, our population is that group of people among whom we are interested in noting behavior changes resulting from, or likely to result from, our educational and organizational methods.

The educational or organizational objectives will determine which people are to be included in the population for a particular study. An objective aimed at finding out what farmers learned about dairy herd improvement practices from our teaching efforts should be evaluated by surveying a population consisting of those farmers who have dairy herds, or expect to have dairy herds, and who were exposed to the learning situation.

An objective dealing with teaching homemakers better child nutrition practices should be evaluated in terms of a population consisting of those women who have, or expect to have, children to feed. Again, some criterion of exposure must be included. An organizational objective to set up a county extension council of local 4-H leaders has a potential population of all the local 4-H leaders in the county.

The population should be identified at the time the plan of work, or the organizational plan, is made up. If it was not done then, it must be done before the sample can be drawn.

The educational or organizational methods used in attempting to accomplish the objectives usually limit the population to a select group within the specialized population, as well as within the total population or the area. Although it is desirable to reach all dairy farmers, or all homemakers with children, or all local 4-H leaders, it is often impossible to do so. Only those who are **reachable** by means of the methods used should be included in a population we plan to sample to provide evidence for evaluation. 4-H Club members make up the population being reached through 4-H Club teaching methods. Only the dairy project members are reached through dairy project work.

Before some of you take exception to the last two statements, we will pause to consider why they were stated so positively. For our immediate purposes, we have to ignore secondary exposure. It is always possible that there is a larger population than the one which we can define in terms of our objectives and those most likely to be directly associated in the learning experience. This larger population may include friends and neighbors of the defined population; it can include other members of the family and casual passers by.

However, it is so difficult to draw this larger boundary that we are forced to ignore it for purposes of evaluation. If attempted, errors of definition could result in defective analysis and erroneous conclusions. If we feel we must know the identity of this larger population we can set up a separate survey or study for that purpose. It is much better to tolerate the delay and drain on resources to make such a survey than to hazard a guess based on the sketchy data we may have at hand.

In your program planning you decide whom you would **like** to reach or teach. By the methods you decide to use, you identify those you **should** reach or teach. Then, **what** you do as a teacher or organizer, and **how** you do it, determine the population from whom you can collect evidence for evaluation. **Who** is in the population, and **where** they are, will determine what methods you must employ to collect the evidence you will need to evaluate your program.

The size of the population will help us decide whether or not a sample should be used. If it is practical to contact the entire population, there is no need to resort to a sample, which always introduces some element of error, regardless of how well it is handled. When we deal with the whole population we do not have to concern ourselves with matters of testing for the statistical significance of our results, or the representativeness of our sample.

This suggests the next term we should discuss, **representativeness.** This is the essential ingredient in the last half of the definition we suggested at the start of this chapter. It requires that the sample from whom we get our information must be representative of the whole population. If our sample is to serve as a reasonably accurate estimate of the group, it must possess all the characteristics likely to be encountered in the group. The degree to which it fails to meet this criterion will be the degree to which it is likely to fail as a spokesman for the group, and the degree to which our conclusions based on the sample data may be in error.

The way you select your sample will determine to a large degree the reliability of your evaluation or study. A device, method, or technique used in sampling is reliable only to the degree to which the results you get by its use approximate true conditions. You may be convinced that at least some of the people in your county actually were changed by your teaching effort. If your sampling techniques are sound, the percentage of people in your sample that changed practices should be very near the true percentage of all the people who changed practices.

We can do two things which will contribute to the accuracy with which our sample approximates true conditions: (1) We can see that a sample of appropriate size is drawn; and (2) we can arrange for appropriate precautions to assure adequate representativeness.

Sample Size

A common error in selecting samples is to take refuge in sheer size of the sample, as though sample size by itself can assure representativeness. Actually, size alone is no proof of goodness or representativeness. If the sample is not representative to begin with, size does not help at all. A sample of 1,000, improperly selected, can be less representative, and can provide less reliable data, than one of 100, properly selected.

The size of the population is only partially indicative of the proper size for the sample. The important factor for deciding on size of the sample is the presumed homogeniety of the people within the population. It is obvious, then, that describing sample size as a certain percentage or proportion of the population is not a realistic clue to adequacy of sample size. In some situations, a 5-percent sample of a small population may be too small; in others, a 5-percent sample of a much larger population may be adequate, or even unnecessarily large.

Proper sample size for the particular evaluation or study depends on: (1) Desired level of statistical accuracy, (2) costs in material and other resources, (3) homogeniety of the population, and (4) contemplated intensity of analysis.

Whenever a sample of the population rather than the whole population is studied, the results have what we call an "error," "standard error," or "probable error." When used as we are using it here, this "error" does not imply that a mistake has been made. Rather, it serves as an estimate of the probable limits of difference between any result we get and the true result. The larger the sample we use, the smaller this error factor becomes.

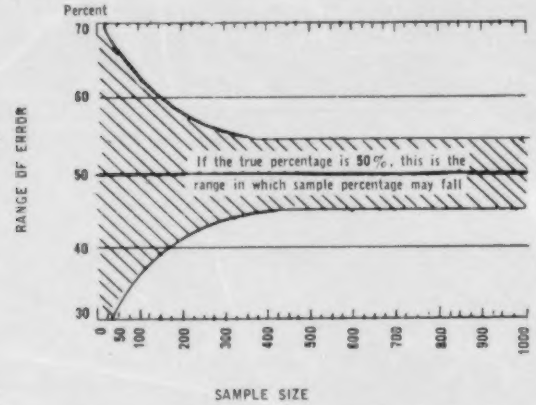


Figure 1.—Diagram showing range within which sample percentage may fall if true percentage from total population is 50—by sample size.

In figure 1, the shaded area shows the range within which a percentage obtained through study of a sample is most likely to fall if the true percentage for the population from which the sample is drawn is 50. If the sample size is 25, and the true percentage is 50, the percentage obtained from a sample might be anything from 30 to 70. In other words, the error might be as great as 20 percentage points in either direction from the true figure.

With a sample of 100, the percentage obtained from a sample might fall between 40 and 60 (a possible error of 10 percentage points). If the sample is increased to 500, the percentage obtained almost always would be expected to fall within 4.5 percentage points of the true percentage.

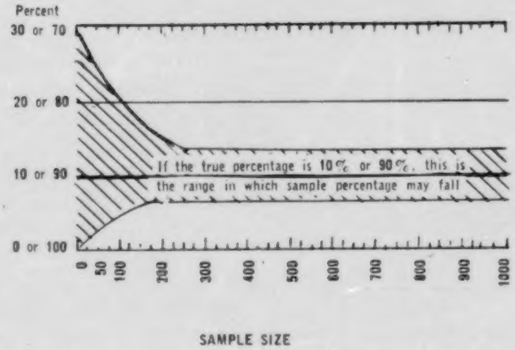


Figure 2.—Diagram showing range within which sample percentage may fall if true percentage from total population is either 90 or 10—by sample size.

The nearer the true population percentage comes to the extremes, 100 or 0 percent, the narrower becomes the range of probable error in the percentage obtained from a sample of the population. If sample size is 25, and the true percentage is 90, the percentage obtained from the sample will most likely fall between 78 and 100.

If the true percentage is 10, the percentage obtained from a sample will most likely fall between 0 and 22. This means that the error might be as large as 12 percent. Referring back to figure 1, it will be noted that for the same size sample (25), the possible error is 20 percentage points when the true percentage figure is 50.

We can conclude that the nearer the population percentage is to 50, the nearer the probable error range will be to that depicted in figure 1; the nearer the population percentage is to either extreme, the nearer the probable error range will be to that depicted in figure 2. Regardless of the size of the true percentage, however, the larger the sample size is, the smaller will be the probable error. (Observe how the range narrows on each of the figures as larger sample sizes are used.)

This consideration of the probable error factor suggests that we cannot predict exact numbers or percentages for our population as a result of our observations of a sample. However, in much of our work, exact numbers are not vital. We are not so concerned with the exact number or percentage of people who have made some desired change in behavior as we are in learning whether or not a relatively large or a comparatively small amount of change occurred.

The **cost** factor includes time and labor and other resource uses, as well as cash outlay. Since this factor looms so conspicuously in any research or evaluation plan, we must work with the smallest sample we can and still feel confident of getting a usably small margin of statistical error. This demands that we be especially careful to make our sample as statistically representative as possible within our size limitations.

Regardless of resource limitations, the sample must be large enough to provide the opportunity for all characteristics we desire to analyze to be in the sample, or representativeness will be sacrificed and the accuracy of our conclusions will be subject to question. If age, education, farming status, nationality, distance from town, and other selected factors are presumed to have bearing on desired behavior change, then the sample must be adequate to assure a strong likelihood that most of these characteristics will be well enough represented to permit a reasonable amount of subsorting on each of them. This means we would want most of the age groups represented, most of the farming status groups represented, and so on.

Furthermore, we would want reasonable assurance that any special groups we desire to examine more intensively will have a sufficiently large number present so that results of our analyses can be relied upon. A category containing less than 30 items will have such a large probable error factor that the findings related to it will have very little, if any, predictive value. For most of our work in Extension, a homogeneous group of 100 items will carry an error factor small enough to allow for reasonably confident applications of the findings.

If we should desire to sort our category containing 100 items into subcategories (sorting the respondents in the selected age group into several level-of-education, nationality, or farming status categories) we would have to use a larger sample so that the statistical error for each category is within usable limits.

Such decisions regarding the amount of subsorting we plan to do must be made before we draw our sample if we want usable data. We will find that simultaneous sorting on more than two characteristics will usually require a prohibitively large and expensive sample, unless the population is an exceptionally **homogeneous** one. However, we can make tentative analyses using several sorting factors if we hold down the number of categories in a classification (older members—younger members rather than 0-14, 15-19, 20-34, and so on, age groups).

The Representative Sample

Random sampling means selecting a sample in such a way that each item or person in the population being studied will have an equal chance to be selected. A sample chosen by random selection is more likely to have the different charasteristics of the population, and in the same proportions, than any other sample you can select. If a sample chosen by this method also has adequate size, it is more likely than not to resemble the population from which it is selected.

Previously, we have been avoiding the use of "selection" in speaking of sampling. Now that we begin using it, we must make clear that we are using it in a statistical sense. Actually we do not select the sample. What we do is decide on the method of selection and leave final selection to the method.

If the method says we take every fifth name from a list, starting with the second name, that is what we do. We don't start with the first name on the list, or take the sixth or fourth name once in a while just for variety. We do not substitute or deliberately omit a name.

There may be occasions when such things are permitted, but they must be covered in the rules agreed upon before selecting the sample. Whatever the justification for their use, they impair the randomness and representativeness of the sample and should be discouraged, unless there are compensating reasons for shifting from a strictly random sample to a modified random, or purposive, sample.

The primary objective of random sampling is to assure equal opportunities for all units to be included. It is not enough to know that certain biasing factors have not been introduced at the point of sample selection; it is also necessary to know that all parts of the population have a chance of being included.

Our aim in drawing a sample that will serve its purpose in evaluation is to have one that will represent the population from which it is drawn. According to Earl Houseman of the Agricultural Marketing Service, "A **representative** sample is a probability sample such that the sampling errors and biases associated with the method of selection, nonresponse, and estimation are known to be sufficiently low to serve a useful purpose."[6]

Several methods of drawing samples by random selection have been used in extension evaluation work. We will discuss several and mention the known limitations they place on true representativeness, along with the reasons which help justify their use.

List Sampling—One method of sampling is to select a group of people from a list. Through the use of random selection methods, **every** person on the list will have just as good a chance of being selected as another. The list must, of course, contain names of the total population in which we are interested. This method should be used whenever a complete list of the total population is available.

There are two favorite ways of selecting a random sample from a list of names. For either method, the names should be arrayed in some manner and numbered. It makes no difference how they are arrayed before they are numbered, such as alphabetically, geographically, by mailing addresses, or by clubs.

The first, and easiest, way is to select every "nth" name from the list. If you need one-fourth of the names on the list, you select every fourth name; if you need one-twentieth of the names to make an adequate sample, you select every twentieth name.

To calculate the proportion of the names that you will require (to determine what the "n" will be), divide the number in your population by the number you desire for your sample. Convert this to a fraction with a "1" for the numerator. Example: There are 7,697 names on the list. A sample of 350 is desired. You get approximately 22 when you divide 7,697 by 350. Therefore, your fraction is 1/22 and you will need to select every 22nd name.

To determine your starting point, select a number by some random method from 1 through 22. Use this number as the first one in your sample and take every 22nd number from that point to the end of the list.

This method meets the criterion of equal chance as long as the original number is selected in a purely random fashion. We can resort to a table of random numbers to determine the first number, or we can place the numbers on separate cards or slips of paper and draw one from a hat.

Sometimes we can pick this starting number in some way which might violate slightly our representativeness ideal, but which will have a symbolic implication for the study. We might use the day of the month, the month of the year, or some other date-based starting point related to the starting date of the study. We might use the hour or minute at the time the starting number was picked.

Suppose we sit down at 10:45 in the morning with our completed list before us and with our sampling fraction calculated. Rather than fumble around with preparing slips for a drawing, we might get on with our work by arbitrarily selecting "10" as our starting number (45 would be too large for our purposes). Then the 10th name would be the first in our sample, the 32nd (10 plus 22) would be the second, and so on.

Another method of random selection from a list is using a **random numbers table.** These numbers have been assembled into tables by purely random means. At the time they were put up in tables, each number had an equal chance of falling in any position in the tables.

To use such a table, you start anywhere on it and take each number which falls within the range of numbers in your population until you have enough for your sample. In our example, with a population of 7,697, we would use the first four columns in each row, because of our 4-digit population total, and include every number less than 7,698 that we come to until we have 350 numbers. (For more information about using random numbers tables, see reference 9.)

For long lists of names, this method is very time consuming. However, it is a more reliable method than selecting every "nth" name if there is likely to be any cyclic or patterning effect in the original listing. Conceivably, you could wind up with all males or all females if you were to draw every "nth" name for a 25- or 50-percent sample of a membership list of a young couple's home planning club.

Area Sampling—More often than not, you will find it necessary to work without a complete list of the population. In order to sample a population whose individuals are not readily identifiable, the system called "area sampling" has been devised. In **area sampling** you choose a random sample of geographic areas instead of a random sampling of individuals. You base your analysis on data secured from all the individuals who live within the selected areas, and who qualify according to your definition for inclusion in the population.

To set up an area sample, a map of the area concerned is marked off into segments which have in them, on the average, a predetermined number of units (farms, households, dwelling units). Whenever possible, boundaries that are easily seen while driving along in an automobile (township lines, railroad tracks, highways, streams) should be used to bound the sampling areas. When an area is so thickly settled that it is impossible to use easily identifiable boundaries, aerial photographs which reveal more minute details of the landscape (fields, farmyards, gardens, orchards) can be used.

Before starting to mark off sampling areas, you need to decide the number of units you wish to have in each area. This number may be of any convenient size, depending on situational and study requirements (homogeniety of the population, desired statistical accuracy of results, contemplated detail of analysis, convenience in interviewing).

Next, you start in the upper right-hand corner of the map and count off a number of units around which you can spot identifiable boundaries. If the number is over half as large as, or is less than half again as large as, the desired average number, it can be considered a sampling area. If the number is much smaller than the desired average, you combine it with the next one you outline. If the number is much larger than the desired average, you divide it through some means other than identifiable boundaries.

Move across the map to the left-hand side, marking off areas as you come to them. Then move back across the map to the right-hand edge. Continue this process until the complete map is marked off into sampling areas. Next, number these areas consecutively, again starting in the upper right-hand corner of the map and moving to the left, then toward the right, and so on, in a serpentine fashion. When all areas are numbered, you sample from them just like you would from a numbered list of names. Sampling can be carried out by either method described under "List Sampling," although the first method is most often used.

In those parts of the country where the grid system of surveying is used in the rural areas, the section and township lines provide easily identifiable boundaries for area sampling segments. In rural areas in other parts of the country, and in the more densely populated areas, plat maps and aerial photographs are used to advantage. You will find that the major problem at this point is getting materials which were prepared recently enough so that you do not have to take time to bring them up to date before you can use them.

An important advantage in using area sampling is that your respondents are concentrated into groups by the manner of selection. This holds down costs associated with travel, maintenance, and supervision of field personnel during the study.

A disadvantage is that, in order to hold the statistical error down to a usable point, a larger total number of people will need to be included in the sample. The geographic areas themselves are considered the sampling units for determining statistical error. For example, if 65 geographic areas which include 325 respondents are covered in a survey, 65 is considered the sample size, and not the 325.

Grid or Random Point Sampling — Another method of sampling geographic areas is to use a "spot" method. One way is to lay a grid over your map and choose the units located at or nearest to the points where the lines cross. You might also use the segments marked off by the grid as areas, number them, and select from among them by using a table of random numbers or some more arbitrary way. You will run into a problem of identification here, however, as you will with any system which does not make use of conspicuous natural or man-made landmarks.

Yet another way is to put a number of dots, hit or miss, over a piece of transparent paper and lay it over the map. You can use these points as centers of sample areas of predetermined size, or you can take the "n" units living closest to this point to be your respondents.

Each of these methods can be deficient in representativeness if the population is likely to be spotty in the way in which characteristics are distributed throughout the area. However, there is a known probability ratio for being included at each stage of the screening process, so they qualify on strictly statistical grounds.

Group Sampling—Sampling groups of people instead of individuals can be done by using the same principles that are used in area sampling. In extension files we have all kinds of lists by clubs, both membership and mailing lists. We can use the mailing lists to classify people into the geographic areas (such as townships) in which they live. Then we can interview all the people who live in designated areas, instead of the entire membership.

If this still leaves us with too large a sample to question, we can sample within the areas by reassembling the lists by areas and drawing a fractional sample from each in the manner suggested earlier for straight list sampling. This is an example of **double-stage** sampling.[10] Other stages can be set up, and often are, but they add immensely to the complexity of the study and should be held down as much as possible.

Using the group sampling technique, we can study club effectiveness by studying representative clubs, or club members located in representative areas of the county or State. The technique works equally well for checking on acceptance rates of recommended farm practices and for observing most kinds of behavior change in relation to the extension teaching practices used.

Unfortunately, this technique increases statistical error, but it decreases cost. If we have reason to believe that differences are as great within groups as between them or, conversely, that one group is apt to be very much like another in regard to the matter under consideration, we can use larger groups for sampling units, thus saving even more time and other resources.

As with area sampling, it costs more when a sample of 5 individuals is chosen at random from each of 20 groups chosen at random (yielding a sample of 100). However, the resulting sample provides more statistically accurate data than does a sample of 10 individuals chosen at random from each of 10 groups chosen at random (also yielding a sample of 100).

Stratified Sampling—Thus far we have described sampling methods aimed at bringing into the sample all characteristics of the people in the same proportions as in the total population. However, if the sample is small, it is possible that characteristics belonging to small groups within the population may be omitted. In order to be sure that certain especially important characteristics get included in the sample, some method of stratification is recommended.

Stratified sampling involves dividing the total population into several separate populations in such a way that these populations are different in regard to selected known characteristics (age, farming status, size of city). If the purpose of using a stratified sample is to assure that certain characteristics are included in the sample in proper proportion, the same sampling fraction is used for each stratum (each separate population).

If the purpose of using a stratified sample is to obtain records from enough individuals having specified degrees of certain characteristics so that these subtypes can be studied separately, a different sampling fraction can be used for each stratum. Whereas one-twentieth of one of the strata may provide an adequate size sample, it may require one-third of another stratum to provide the same size sample.

This technique is used very frequently in studies involving farm operators. We realize how widely farm operations vary in size, and we also realize that behavior of farm operators on the different sized holdings is likely to differ as widely. To be sure of adequate numbers of respondents with operations in the different size categories, we set up several different sampling fractions. We might have to use all the very largest operators in the area (1/1 fraction). A 1/10 fraction might do for operators with holdings in the middle range. Among the much more numerous small or subsistence operators, a 1/100 fraction might be adequate.

When this method is used, data from the different strata cannot be assembled by simple addition for purposes of obtaining totals for the population. Data from each stratum first must be weighted according to sampling fraction used. In our farm operator sample, we would have to invert the fraction for each stratum, convert it to a whole number, and multiply by this term, before adding up to get population totals. We would count our large operator data as observed, but we would have to multiply our middle range operator data by 10 and our small operator data by 100 before totalling.

Judgment Sampling

The sampling methods we have been describing depend on objective procedures which should result in fairly representative samples (everyone in the population has an equal or known chance of being selected). There are occasions, however, when you might find it better for your purposes to choose a sample on the basis of what you believe it should contain. In so doing, you hand-pick each individual to go into the sample, deliberately making up what you feel is a sample representative of the population from which it is chosen.

Yates, the English statistical authority, has this to say about this method:

"The ideal which is aimed at in sampling is to make the sample as representative as possible, so that measurements or observations on it can be taken as virtually equivalent to similar measurements on the whole population. The fact that this ideal is in the mind of the sampler when taking the sample naturally influences his selection if he has any freedom of choice. Most samplers when selecting a representative sample will deliberately reject abnormal material, or if they feel that the sample should be representative of the abnormal as well as the normal will deliberately balance up the different categories abnormally.

"Unfortunately the sampler's claims to be able to select 'a representative sample' by personal judgment are largely unfounded, and his selection is in fact subject to all sorts of biases, psychological and physical. To avoid these biases and to provide an estimate of the representativeness of the sample, i.e., of the 'sampling error', more rigorous processes of selection has been devised."[12]

Houseman comments: "It is reasonable to expect that judgment samples will usually be selective in the direction of what the mind regards as the more important elements or in the direction of the less obscure elements. It is also reasonable to expect in general that the variability would be less within judgment samples than within the whole population; if so, this would mean that distributions based on judgment samples might be seriously distorted . . . Judgment samples naturally look good to the person who made the selection, but since a sample cannot be appraised by its looks, judgment samplers are likely to have erroneous impressions of the accuracy of their samples."[5]

Limiting the Population to Study Size

Within certain populations, especially when variability is great, it may be advisable to define the population in such detail that all or nearly all of the members of the population which have the desired characteristics must be included in the survey. In such cases, sampling, as defined in this chapter, is neither necessary nor appropriate. If you should wish to study 14-year-old, 1st-year 4-H Club members in a county, you would seldom find enough of them to make up a population large enough to sample. You would have to use all of them.

We must remind ourselves constantly that, as a result of narrowing the population to study size, we can apply the findings only to the group we have studied. Comparisons with similar groups of 14-year-old, 1st-year 4-H Club members in other counties may be made, but only with due allowances for differences in environmental elements of the different locales. We can draw no direct inferences to another or larger group or population. Anyone who desires to use such data for purposes of comparison must be made aware of these limitations, and should use them, if at all, only after careful consideration and interpretation.

Testing the Sample

We have statistical formulas for determining whether or not a sample is of adequate size to accomplish the objective for which it is drawn. We also have checks on representativeness. We can use the following methods quite effectively in extension evaluation.

If the population to be studied is a class of people for whom information is available from census sources, we can check data from the study against comparable census data. Sometimes, when information requested for census purposes may not be relevant to our particular evaluative effort, we may ask a few census-type questions, anyway, so that we can have data to check with census materials to get a line on the representativeness of our sample.

Another test for adequacy is to draw from the population several other samples of approximately the same size as the given sample. Then we can compare the results for the different samples. Unfortunately, this can become prohibitively expensive. A more economical technique with the same objective is to divide the sample into two parts and compare the results.

However, even when such are used, they do not serve as a complete check on representativeness. Of greatest importance is the assurance that extreme care has been taken in following the sample selection method, in obtaining information from the total sample, and in collecting information in such a way that biases and errors are minimized.

Summary

Most evaluation problems require sampling in some form. For that reason, we have devoted this chapter in our methodology section to a discussion of what sampling is and how it works.

After some basic definitions and a rundown of the four steps involved in a typical sampling problem, we discussed several ways to sample and mentioned some of the conveniences and shortcomings of each. We included some of the more approved methods, as well as several which may have to be used when conditions will not permit an entirely objective sampling plan. We stressed the importance of size and representativeness in a sample and suggested some ways to check on each.

We discussed how probability or chance enters into the rationale of inferring from a sample to a population. Without going into the mechanics of statistical computation, we commented on the importance of using statistical tests for probable error and reliability.

References

1. Croxton, F. E. and Cowden, D. J. **Applied General Statistics.** New York, Prentice-Hall, Inc., 1955.
2. Gallup, George. **A Guide to Public Opinion Polls.** 2 ed., Princeton, N. Y., Princeton University Press, 1948.
3. Garrett, H. E. **Statistics in Psychology and Education.** New York, Longmans, Green & Co., 1939.
4. Hartung, M. L. and Taeuber, C. Talks by. **Report of University of Chicago Evaluation Workshop, 1946.** U. S. Dept. Agr., Ext. Serv. Stencil No. 781 (5-46), Washington, D. C., 1946.
5. Houseman, E. E. **Designs of Samples for Surveys.** Agr. Econ. Res., v. 1, No. 1, January 1949. (Published by the Bureau of Agricultural Economics, U. S. Department of Agriculture.)
6. Houseman, E. E. **Talks with Him in Connection with This Chapter, with Earlier Papers on Sampling, and with Sampling Problems Arising in Our Division's Work.**
7. King, A. J. and Jesson, R. J. **The Master Sample of Agriculture.** Jour. of the Amer. Statistical Assoc., v. 40, March 1945.
8. Pearson, K., edited by. **Tracts for Computors.** No. XV. "Random Sampling Numbers." Arranged by L. H. C. Tippett. Cambridge University Press, 32 E. 37th Street, New York 22, N. Y., 1927.
9. Sabrosky, L. K. **Instructions for the Use of Random Sample Numbers.** U. S. Fed. Ext. Serv., Mimeog. 206 (3-49), Washington, D. C., 1949.
10. Sabrosky, L. K. **Percentages Obtained from Samples as Compared with Percentages Obtained from Total Population.** U. S. Fed. Ext. Serv., Mimeog. 205 (3-49), Washington, D. C., 1949.
11. Yates, F. **Sampling Methods for Censuses and Surveys.** New York, Hafner Publishing Co., 1949.
12. Yates, F. **Some Examples of Biased Sampling.** Ann. Eugenics, 6:202-213, 1935.

Chapter XI

METHODS OF COLLECTING DATA

Gladys Gallup

Up to this point we have been discussing the objectives of the evaluation, evidence of the objectives, and the persons or population from whom or about whom we will collect the evidence. In this chapter we will discuss methods for collecting the evidence or data.

The method of collecting the data should never determine what we collect. The most important point is to be sure of exactly what information we need to meet the requirements of the study. The person conducting the evaluation then needs to decide what is the most practical way to collect the data. He needs to consider the type of device he will need to get a record of evidence. He will also also need to consider the conditions under which he will have to work, and other demands on his time, budget and other resources.

Methods of Collecting Data

1. Mail questionnaires.
2. Personal-interview schedules, usually used while interviewing people in their homes.
3. Distributed questionnaires and checklists of all kinds.
4. Group interviews.
5. Case studies.
6. Systematic observation procedures.
7. Systematic studies of available records.
8. Tape recorded interviews, discussions or panels.

Considerations Which Influence Selection of Method

1. The objectives of the study.
2. The type of device that might be used to obtain evidence.
3. The respondents from whom you will need to get the data.
    a. Where and how can they best be reached—at home, through the mail, at meetings, or elsewhere?
    b. Will they represent a sample of the total population which the extension worker has attempted to reach?
4. The resources you have available.
    a. Your time—demands of other work.
    b. The time of others who will also work on the study.
    c. The time it will require of the respondent, the person who is to answer the questions or give the information.
    d. Money available for supplies, stenographic and statistical help, travel, subsistence.
5. The advantage and limitations of each method.

Analysis of Each Method

In selecting the method to use in collecting data, the evaluator needs to analyze each method to see whether or not it will provide him with valid and reliable information. Each method has its advantages and limitations. These should be taken into consideration in making decisions. Let us first take up the mail questionnaire.

The Mail Questionnaire—This is sent to respondents who fill it out and return it by mail. Before deciding to use this method, there are certain questions which we need to consider.

1. Can the questions on the objectives of study be fitted to a mail questionnaire?
2. Can the questions be stated in a clear and understandable fashion without further explanation?
3. What about the people returning them; are they representative of the group with whom we have been working or of the "population"?
4. What about response rates?

Advantages of the mail questionnaire:

1. It is an easy, quick and relatively inexpensive method of obtaining data, especially when compared to a personal interview designed to secure the same information.
2. It can be planned to reach a wide geographic distribution of people including some quite inaccessible people, provided that their addresses are known.
3. It is useful in reaching a relatively homogeneous, fairly well-educated group.
4. It is well adapted for reaching specific groups such as dairy farmers, poultrymen, or homemakers who belong to organized home demonstration groups.

5. It can be free of any interviewer bias, but not necessarily instrument bias.

Limitations of the mail questionnaire:

1. The questionnaire must be short.
2. It is difficult to obtain detailed qualitative answers or to know precisely what the responses mean.
3. There is little control over completeness of answers; when the respondents don't understand or don't want to say, they may just leave out the answers to a question.
4. Those who reply may not be typical members of the list; those who have made changes in practices or who are especially interested or who are especially opposed to the ideas presented are the more likely to reply.
5. Questionnaires should be returned from everyone in the sample. This usually requires followup. It takes time to write followup letters, to make telephone calls and to call up respondents who fail to reply.

Personal Interview Schedules—These schedules are filled in by an enumerator or reporter. This method is used in a more formal type of study in which an interviewer calls on a sample of people. The interviewers are usually neighboring county extension workers, State extension workers, and lay people, such as extension committee members.

Advantages of personal interview method:

1. The people who participate have an opportunity to observe and study situations and conditions; they talk directly with people and get their reactions to practices and programs.
2. The personal interview method usually yields a high percentage of returns, as most people are willing to cooperate.
3. The interviewer has an opportunity to explain questions to respondent.
4. Complete answers to all questions can usually be obtained. This contributes to statistical accuracy, validity and reliability.

Limitations of the personal interview method:

1. Transportation costs and time required in a large area may make the personal interview method impracticable. For example, good enumerators, unless volunteers, can demand very good salaries. An enumerator might be tied up several hours on a single interview.
2. Unless the interviewers are properly trained and supervised, the data recorded may be inaccurate and incomplete. A few poor enumerators can bias the study.
3. It is usually claimed that costs per interview are higher than when mail questionnaires are used. This may not be true if the area to be covered is not too great. The personal follow-up of the mail questionnaire in order to secure records needed may run the cost as high as for personal interviews.

Distributed Questionnaires or Checklists—These instruments are often used interchangeably. They are usually handed directly to the respondents individually or in groups. Respondents answer the questions and hand them back to the person conducting the evaluation.

Respondents are asked to record their answers to questions by checking or otherwise indicating their selection from a list of statements. These are usually distributed:

1. To all persons in a group or in attendance at a meeting, to be filled out at the meeting.
2. To a sample of people who are participants in a group or meeting, but to be filled out at home.
3. To leaders who ask their neighbors or members of the group for information.

These small report forms are used extensively in home demonstration work. They are used to collect evidence of progress made—usually practices adopted—for use in program planning and in writing the annual report.

Advantages of the distributed questionnaire:

1. The questions have a definite relation to recognized goals or objectives.
2. The cost is small and they are easy to administer.
3. They are easy to complete.
4. They provide a quick way of surveying the group.
5. When used before and after a meeting or a series of meetings, they can be used as interest getters.
6. If filled out in a meeting, they can be used as a basis for discussion and can indicate the next steps in program planning. (People involved in filling out the forms are also usually involved in program planning process.)
7. They can also be used as attention holders and indicators of action.

Limitations of the distributed questionnaire:

1. The people who fill out these forms are usually participants in the program and know the objectives and may be inclined not to be critical of the existing program.
2. These report forms may center attention on "practices adopted" rather than on other kinds of behavioral changes.
3. Participants in groups in an area are not usually representative of all people in the area. Therefore, it is better to interview a

cross section of people in the area if we desire to measure changes made by others than those in the organized group. In other words, these individually distributed questionnaires tend to measure progress only of these who participate in the program.

4. They are usually on single topics and therefore cover only a limited field of study.

Group Interviews—The leader of the study presents a questionnaire to a group of people assembled in one place. Each person in the group is asked to record his or her answers to the question as it is read. To secure an adequate cross section, those people interviewed must be carefully selected so as to secure representative samples of the people about which generalizations are to be made.

Advantages of group interviews: (Advantages are similar to those of the individually distributed questionnaires.)

1. If the persons in the group are homogeneous and if the persons attending the meeting are representative, data are easily and readily obtained.
2. There is very little cost—easy to administer, very little travel, a small crew can do the job.
3. Speedy—the whole interview can be conducted in a very brief time.
4. There will be few refusals.
5. The questions can be used as a basis for discussion and can help toward program planning, as well as to hold or get interest.

Limitations of group interviews: (Limitations are similar to those of the individually distributed questionnaires.)

1. The replies may be unconsciously or otherwise influenced by those with strong opinions (peer group bias). Since you want independent judgments, discussion should be discouraged during the interview.
2. The people attending a particular meeting are seldom representative of the entire group.
3. The questionnaire used for group interviewing is usually short and therefore can cover only a limited field of study.

Case Studies—These can be studies of one or a few individuals or cases, a chronological report of the development of a project or activity, or a story of an individual family. Frequently they are reports of concrete events which are vivid, convincing stories of the effects of a program. They may constitute important evidence, if properly assembled.

Advantages of case studies:

1. They give concreteness and reality to problems and solutions.
2. They present the situation, reveal the importance of human factors in organization, show the difficulty of problems as they evolve.
3. Often a case study will bring out important factors which are assumed to be general knowledge and for that reason go unmentioned by respondents interviewed by other methods.
4. They may show how decisions are made.
5. They show the importance of timing.
6. They show the continuity of development and how the project or activity unfolds.
7. They bring out the problems as they arise, how these problems were solved, and how the mistakes were made.
8. They afford an opportunity to study a process minutely and to discover its strengths and weaknesses.
9. They are effective in serving as springboards to further thought and consideration—a good teaching device.
10. They show Extension's contribution in helping to solve problems.

Limitations of case studies:

1. Case studies may give particularized data, they may emphasize the unique.
2. They may provide data on only 1 or 2 aspects of a problem, and not on all relevant aspects.
3. A case study is only one case; unwarranted generalizations may be made.
4. They are time consuming and require training and experience in observing, recording and writing.

Systematic Observation Procedure—This method is used during visits to homes or farms. Observation of practices which have been recommended for the area are observed and recorded in some systematic manner.

Advantages of systematic observation:

1. The worker knows the recommended practices.
2. He can observe to what extent they are being adopted.
3. He can have them listed on an observation sheet and check them off. This avoids hit or miss observation.
4. He gets firsthand information.
5. He can obtain permanent record information at the same time to supplement his observations.
6. Progress can be shown.

Limitations of systematic observation:

1. The observer may use his own values—attribute importance to what he thinks is significant. An objective check sheet helps to prevent this.

2. It is difficult to separate observation from interpretation.
3. Enumerator has only his intuition to tell him what is representative.
4. Costly when much information must be collected.
5. Impractical to do for much information.
6. Details important for analysis may be omitted.

Systematic Study of Available Records—4-H Club project records and farm and home development records are frequent sources for this kind of analysis. An analysis of records and reports involves much more than recording, tabulating and reporting.

Advantages of records analysis:

1. Data can be obtained directly from the records without going through a field survey or some form of interviewing.
2. This may take the time of just one person.
3. They are reliable if the records have been carefully kept.
4. They provide information from the past which could not be collected today.

Limitations of records analysis:

1. Records may be carelessly kept and incomplete.
2. Records may not be from a representative sample of the population.
3. Often there is no opportunity to make checks on accuracy of entries.
4. Records may lack comparability and, therefore, be difficult to analyze.

Planning to Carry Out the Evaluation

We have discussed seven different methods of collecting data. Deciding on the data collecting method does not end our decision making action, however. We need to decide not only on the method of collecting data but also **who** will be responsible for gathering the information, **when** the work will be done, and **where** it will be done.

Who Will Be Involved in Gathering the Information?—This varies with the type of information wanted and the method used in collecting data. If a mail questionnaire is to be used, only the county agent may be involved. However, lay people should be involved in planning the evaluation and whenever possible should be used in collecting information.

An important value resulting from a study is the training and experience gained by participants. The people who are involved in the development and execution of a program should be involved in its evaluation. These may be:

1. Agent alone, or with other agents from home county.
2. Agents from another county.
3. Agents and lay people.
4. Agents, lay people, and State staff personnel.
5. Agents, lay people, State staff, and Federal staff personnel.

When Will the Information Be Collected?—It is desirable to decide considerably in advance the exact period that will be covered by the study—the month or season, a specified time after the work has been taught or the activity completed. The evaluation should be planned so as to avoid heavy work seasons of farm people.

When the information is to be collected depends also on the other jobs that must be done in the evaluation and the time that must be allowed for each job. Very early in our planning we should work out a detailed schedule spelling out, (1) jobs to be done, (2) when each job is to be done, and (3) who is responsible.

Where Will the Study Be Made?—Step-by-step development of plans have been covered up to this point. The method used in collecting data, and from what people, will determine where the study is to be made.

Summary

The methods of collecting data should not determine what to collect. Information or evidence needed to meet the requirements of the study is of first importance. The method used to obtain evidence depends on the kind of evidence you need and on the resources you have to collect it.

Some methods of collecting data are: Mail questionnaires, personal-interview schedules, check lists of all kinds, group interviews, case studies, systematic observation procedures and studies of available records.

References

1. Barr, A.; Davis, A.; Johnson, P. O. **Educational Research and Appraisals.** New York, J. B. Lippincott Co., 1953.

2. Parten, M. B. **Surveys, Polls, and Samples.** New York, Harper and Bros., 1950.

3. Riecken, H. W. **The Volunteer Work Camp: A Psychological Evaluation.** Cambridge, Mass., Wesley Press, 1952.

4. Sabrosky, L. K. **Discussions — Evaluation Workshop,** April 5-23, 1954, Aguas Buenas, Puerto Rico. USDA, Ext. Ser. mimeograph ER&T 1-(1-59). Discussion VII—How You Will Get the Information.

Chapter XII

INTERVIEWING

Mrs. Jewell G. Fessenden

Throughout the process of developing a study plan we have to keep in mind how the information will be collected. In this chapter we will discuss interviewing as a method of collecting data. A good interviewing job cannot salvage a poorly designed study but poor interviewing can lessen the value of a well-planned study.

We need to know what kind of interviewers we will have before we can decide how involved to make our plans for collecting data.

We must know whether interviewers will serve without pay or will expect to be paid, and how much, before we can make estimates of costs for the study. We must know how many we will use before we can estimate how long the data collecting phase is likely to take, and whether or not we will be able to handle supervisory details without assistance.

An **interview** is a contact by an investigator with a person for the purpose of securing specific information. It can take place in a face-to-face setting with the interviewer asking questions and recording the answers he receives from the respondent. It can occur in a modified face-to-face setting where the interviewer asks for information from a group of people. With this arrangement, respondents will record their answers in some manner and the interviewer will be concerned, primarily, with clarifying statements and striving for completeness of response.

A third type of contact is by means of the telephone, wherein the interviewer solicits answers from the respondent via the telephone and records them on some sort of tally sheet or schedule.

Most of us in Extension do a good amount of interviewing at one or more of these levels. Often it is so casual that we do not think of it as interviewing, and may give no serious thought to evaluating the answers. Occasionally, though, we spot-check some phase of our work by interviewing members of our clientele, and once in a while we enter upon full blown evaluation problems requiring organized interviewing plans.

In addition to supplying us with the information we need for our work, personal interviewing gives us a better understanding of the people with whom we work and of the community in which our work lies. This is most true of the more personalized forms, but is proportionately so for all forms of interviewing. All too frequently, however, we do not give much thought about how to do it properly, or to recognize all the occasions when interviewing may be used to advantage.

Major Kinds of Interviews

Proper collection of data by means of interviews is a very important tool in research.

As we discovered while discussing other techniques used in evaluation, those we can use for interviewing also are varied, and equally subject to complicating factors and limitations. We will talk about the three major kinds of interviewing that can be used to advantage in Extension and mention some of their advantages and disadvantages.

Face-to-Face Interview—This is the most personal level of interviewing and, because of the one-interviewer-one-respondent pattern, it is the most expensive. It is more time consuming to make the individual home visits. Although there are occasions when it is possible to get the respondent to come to your office to be interviewed, most of the time you have to go to him.

Original transportation costs are greater per respondent, and several trips may be necessary before the prospective respondent is contacted and the interview is secured. There is also the strain of repeating the processes of gaining admittance and establishing rapport (making the respondent want to cooperate).

To compensate for the heavy drain on resources required by personal interviews, there are several convincing advantages. The interviewer may be able to keep his respondent interested and attentive for a longer period of time by changing the pace or taking a conversational break whenever he observes signs of declining enthusiasm. This is difficult to do in a group situation or during a telephone interview.

A personal visit provides the interviewer with a chance to observe personal reactions which can be much more enlightening than the vocalized answer alone. He is in a better position to note resistance and may be able to discover reasons and deal with them more easily in a personal setting.

Even though some people may respond more readily to a telephone interview, the interviewer has more opportunities to establish friendly relations during a face-to-face visit. A smile, a comment on something visible, a personalized word of praise, an expectant attitude, all can be used by an alert interviewer to draw out an answer from a reluctant respondent.

Telephone Interview—The telephone interview may be very useful for certain purposes. If the interview schedule is relatively brief, we can assemble complete and accurate information in this manner. However, it is not so effective when "thought" questions are employed or when data which may not be readily available are required. In exceptional cases, and when skilled workers are operating, it is possible to keep a respondent interested and attentive for a half hour or more.

If the respondent happens to be a member of a club, or some other group which is sponsoring the study, and already is familiar with the purposes for which the data are being collected, she may respond quite satisfactorily to a telephone interview. In situations such as this, when good rapport previously has been established, telephone interviewing can be quite effective, and the time and money savings become important inducements to decide on this method. In general, though, telephone interviews are not so desirable for long, tedious interviews, regardless of how well rapport has been established. Then, too, the values accruing from personal observations are lost, and it is not so easy to discover and overcome real reasons for reluctance in answering.

Group Interview—This technique is not to be confused with group counseling or group therapy, certainly not with group discussion. Essentially, it is a device for shortening the data-collecting phase to a minimum. In this procedure, the interviewer assembles his prospective respondents into a single group or series of groups, and explains the purposes of the study to them.

He gives necessary instructions and explains questions which may not be clear. Then the individuals fill out the form. They may be asked to answer each question simultaneously, with the interviewer commenting on each as they come to it, or each respondent can work at his own speed and turn in his schedule when he has worked through it to his satisfaction.

In addition to setting up the interview and making initial explanations the interviewer is there to provide whatever help he can to assure complete and usable record forms. If he is interested in having all of the respondents deal with the same question at the same time, he should arrange for assistants to circulate among the group to aid those who are having difficulty with the meaning of questions and to encourage them to answer all questions. The interviewer should discourage conversation which might lead to pooling of answers, which means getting one set of answers repeated several times rather than several different sets of answers.

The obvious advantage of the method lies in the savings of time and transportation expenses. Then, too, the same preliminary explanatory remarks will serve for the entire group. Some of the more obvious trouble spots also can be cleared up collectively. A third factor is that there should be little or no difficulty connected with gaining admittance, establishing rapport, or dealing with refusals.

If the meeting is called for the specific purpose of filling out the schedule, or if the members know the interview is to be an item on the agenda, they will come knowing the purpose of the meeting and be prepared to cooperate. Similar reactions may be expected from a selected sample which has agreed to assemble for purposes of supplying the information you need.

Limitations of the effectiveness of group interviewing are related, (1) to reluctance of respondents to ask about points which are not clear to them, (2) to their overwillingness to skip the difficult (very likely the most important) questions, and (3) to the pooling problem already mentioned. There is less opportunity to probe for reasons for "yes-no" or "agree-disagree" answers, so requests for free response or essay type answers are not productive.

Responsibility of the Interviewer

The interviewer has a decidedly important part in the successful completion of a piece of research. On him falls the responsibility for actually contacting the respondents and collecting the data. A good device for collecting evidence can be made worthless in the hands of an inefficient or irresponsible interviewer. Because his part is so vital in any research project requiring interviewers, his selection and training are important steps in the research process.

Before interviewers are selected you should have a clear picture of what their duties will be and what will be expected of them. The interviewers should also understand their responsibilities and be willing to accept them. This knowledge will help you rule out obvious misfits, and suitable training should make the others able to carry out their assignments in adequate fashion. We can think of several personal and professional responsibilities of our interviewers. We will discuss each briefly.

1. **Maintain a neat personal appearance**—As your interviewers represent both you and the agency which is sponsoring you, it is best that they serve as good representatives. For this reason, you should try to get pleasant appearing, neat persons to serve as interviewers. They should dress appropriately for the season and the environment in which they will be working. They should be clean and not given to offensive habits or mannerisms.

2. **Locate the respondent, get the facts, and record them** — The interviewer must be familiar enough with the area to which he is assigned to locate the proper respondents according to the sampling plan for the study, or he must learn how to use the map of the area. If a person to be interviewed is not available, he should report the fact to his supervisor. Under no circumstances should he make substitutions without authorization.

A prospective respondent is not dropped from the study or substituted for merely because he is not at home at the time of the initial call. If it is determined that he will be away for the duration of the field work period, that he is ill, or that he will not be available for some other reason, it may be necessary to list him as "not available" or to substitute for him according to the prearranged plan. Several calls may be necessary to locate a person or to complete an interview with him if he works away from home or is busy.

An interviewer must make every effort to obtain adequate and objective answers to all questions. Incomplete schedules and partial answers tend to destroy the randomness of a sample and make study findings less useful. The interviewer should have a personal desire to make every contact fruitful. Notations should be legible, and as complete as necessary for others to understand. Researchers can help interviewers by doing a good job of orienting them to the study and by providing them with good instruments (questionnaires, schedules).

3. **Be considerate and honest with respondents**—Because your interviewer must be able to think on his feet and be able to adjust to the situation with little or no foreknowledge of what it may be, it is desirable that he be an intelligent and understanding person. He must deal in an honest and forthright manner with all respondents, but he must know when to say gracefully that he cannot answer a question. However, questions about the purposes of the study should be answered frankly.

4. **Understand the purposes of the study**—The good interviewer realizes that to do his job right and to have a personal feeling of satisfaction in doing it, he must know what the study is about, what its objectives are, and why the particular questions are being asked. When he understands these points thoroughly, he finds it easier to gain an inner conviction that the study is worthwhile, and also finds it easier to transmit that feeling to his associates and respondents. Similarly, he finds that his work becomes more effective and more enjoyable.

5. **Be thoroughly familiar with the instruments**—It is very important that an interviewer study and understand very thoroughly the forms upon which answers are to be recorded. Once he has gained personal understanding, he can transpose the technical language of the project outline into easily understood and meaningful language for his respondents.

He can gain the desired degree of familiarity by careful attention during the training phase and by asking questions to clear up points which he does not understand completely. When he knows the "why's" as well as the "what's" of his assignment, he can do a better job of anticipating questions of respondents and of having answers ready for them.

6. **Follow sampling instructions**—The interviewer must realize that considerable thought is given to identifying the people he is to interview, and that he has no authority to take liberties with the sampling plan or instructions for its use. Once he is assigned certain persons to contact, it becomes his duty to locate those persons and to do all that is in his power to secure the information.

If there are good reasons why he cannot carry out the plan, he should report the fact, and explain why, so that proper corrective measures may be applied. He must realize that he can make the situation worse instead of better by acting on his own initiative rather than checking with his supervisor for further instructions.

7. **Ask questions exactly as written**—The interviewer must realize the difficulty which lies in designing a question which means the same thing to everyone. At the same time he must realize it is not wise to modify questions at any time. Even if the respondent appears to grasp a different meaning than the one he believes a questions contains, he is not at liberty to interpret the question to get what he considers to be the proper answer.

Although the interviewer is encouraged to urge the respondent to think and to answer completely, he is not authorized to **direct** the thinking of the respondent in any way. He nullifies the purpose of the study if he rephrases questions, suggests things to consider, asks leading follow-up questions, or does anything but encourage the respondent to think for himself.

8. **Check work for completeness (editing)** — No record form should be turned in by an interviewer until he has checked it thoroughly for completeness and accuracy. It is most effective when done im-

mediately after the interview, either before leaving the respondent or before going on to the next one. Editing should be very thorough at this stage.

If the respondent seems to be anxious to have the interviewer leave, the interviewer may thank him warmly for his cooperation and make his check in the car. However, he is in a better position to make corrections and completions if he makes the check before leaving. Ordinarily, a remark like "You have been most cooperative. Do you mind if I glance over this form to see that I have everything?" will take the pressure off the respondent and make him happy to let the interviewer read through the form before leaving. It may even prompt him to volunteer more useful information.

Selecting and Training Interviewers

Extension research usually is conducted under the leadership of extension personnel. For the most part, data collecting is done by extension workers and lay leaders. Sometimes, experiment station and other research persons from land-grant institutions are involved. It is rarely the case that professional research organizations are employed to do research for Extension.

Suggestions that follow will, for the most part, be directed toward the untrained interviewer. More than likely, these will be county level extension personnel or local leaders. Extension personnel do not receive extra pay for such tasks and local leaders are ordinarily not paid for their services.

The absence of direct economic motivation suggests two possible limitations on the effectiveness of interviewing by such persons. For one thing, strict attendance and punctuality cannot be demanded of interviewers and, for another, there may be less feeling of responsibility on the part of volunteers. However, we in Extension have been very successful on many occasions in the use of volunteer lay leaders and county extension personnel in our research teams.

Interviewing is best done by people who are interested in doing the job, and are willing to maintain an objective approach. Before agreeing to help with a study, anyone offering his services as an interviewer should feel reasonably certain that he will have enough free time to stay with the job until it is completed.

You should mention the probable duration of the work when you make inquiries about interviewers. You should include time for training and other orientation activities, as well as time for doing the interviews. You should allow no one to participate in data collection for the study who does not participate in the training phase, and who is not willing to assume all proper responsibilities of interviewers.

For most studies, you can work lay leaders into your interviewing plans with little difficulty. You have no housing problem with them and the transportation problem is minimized. They probably know the area quite well and know how to get around in it.

They are more likely to have a personal interest in the outcome of the study and may be in a position to capitalize on the findings. If properly cautioned, they can be counted on to get accurate information. Their familiarity with the local scene can be used to advantage in identifying and locating respondents.

Regardless of experience, all interviewers require a certain amount of training for a specific research undertaking. The inexperienced interviewer will require more training in the general philosophy and discipline of interviewing, but all of them will require orientation for the particular research problem, and experience in dealing with the specific research instruments which will be used.

We recommend that all interviewers be given training in six areas, with the intensity of the training in each area to depend on the experience level of the interviewers for the particular study. The suggested areas are: (1) Background and purposes of the study, (2) sampling procedures and how to apply them, (3) interviewing procedures and techniques to be used, (4) familiarization with the research instruments to be used, (5) experience in filling in and editing the record form, and (6) experience in asking the questions. An actual field test to follow the training session is very effective in making the transition from theory to reality.

Almost invariably the length of time that can be devoted to training the research team is limited. This necessitates careful planning and efficient use of time. A minimum of six hours is recommended for larger studies, although a longer period is desirable.

It is a good idea to divide the training into two parts: (1) **Orientation** (purposes, background and techniques); and (2) **practice** (experience in filling in and editing the record forms and in asking the questions). We have used the following outline on several occasions in a one-day training session for some quite substantial studies. It requires at least six hours but might take more, depending on the amount of discussion and elaboration you permit.

Agenda for Training Interviewers

1. Why we are here. (Resident agent in charge.)
    a. Greetings and introductions.
    b. Purpose of meeting.
    c. Explanation of plans for the day.

   d. Why we wanted to make this study; how we plan to do it. (1) Review of objectives.
2. What an interviewer does. (Training leader takes over.)
   a. Importance of the interviewer in a research team. (1) Importance to success of project.
   b. Definition of terms. (Respondent, sample, schedule, refusal, call-back, unit, and other general terms. Also any special terms used in the particular study.)
   c. The sampling plan; what it is and how to follow it.
   d. Detailed explanation of instruments. (Question by question and with opportunity for questions from interviewers for clarification.)
   e. General instructions for interviewing. (Have a set of instructions to distribute to interviewers but discuss them in detail at this point. Ask for questions on any point which is not clear to all.)
   f. Specific instructions for asking questions. (After a general statement regarding techniques of asking questions, use the questions in the study instrument to demonstrate how to ask them properly. Also demonstrate suitable forms of approaching the respondent and of terminating the interview.)
3. Filling in and editing schedules. (Have interviewers pair off and take turns asking the questions, and answering them. Each then edits the schedule he has taken.)
4. Editing by trainers. (Trainers check schedules turned in during practice and point out errors to the group to alert them for similar situations in their own interviews. Trainers explain how to avoid errors and how to go about correcting them.)
5. Explaining special procedures for the study.
   a. Procedures for identifying respondent, making contact, arranging for interview; dealing with refusals, not-at-homes, call-backs.
   b. Where and when to bring completed schedules.
   c. How to contact leader of survey party for further instructions or help in dealing with unexpected or unusual situations.
6. Making first assignments for interviewers.
7. Setting time for general meeting at end of first day of interviewing to discuss problems, clear up questions, and check on quality of work.
8. Sumary of day's activities; final request for questions.
9. Closing words of encouragement, assurance and morale boosting.

Interviewing Procedures and Techniques

There are four main stages to making and completing an interview.

1. Gaining entrance, making the approach and establishing rapport.
2. Securing and recording information.
3. Closing the interview.
4. Editing.

Establishing Rapport—A good interview depends to a great extent on a quickly established feeling of trust and confidence in the interviewer by the respondent. Sincere friendliness on the part of the interviewer is a firm step in the direction of achieving this objective. Techniques for gaining admittance to a respondent's home and for establishing friendly relations (rapport) vary with the situation. A good interviewer will learn how to make on-the-spot adaptations to fit most situations likely to confront him.

Some techniques may be acquired from study of literature on the subject, but practice is also important. An interviewer needs to try to understand the reasons for resistance and to think of ways to overcome them. The most common reasons for resistance are shyness, fear of being sold something, feeling of inability to answer questions, and reluctance to talk about personal matters.

Neat personal appearance and appropriate dress are important in creating good first impressions. Equally important is the general tone of your introductory remarks. They should (1) identify you as the interviewer, (2) explain briefly the reason for your call, and (3) ask for the respondent's cooperation.

One interviewer has used successfully the following as an opening sentence: "Good morning (good afternoon, hello). I am (name)." Pause. "May I come in and visit with you for a few minutes?" This provides an opportunity for the respondent to give his or her name, if it is not already known to the interviewer.

Most people invite the interviewer in on this remark. However, if there is hesitancy or if some question is asked as to the reason why, the next step is to explain in a little more detail what you are doing and why. It may help to include some statement to let the respondent know that cooperation is important. "You are one of a few people selected and your opinion (cooperation) is very important," is helpful.

"How was I (my name) selected?" is a question often asked by a person being interviewed. The answer to this should be honest and brief. Something like, "Well, we can't visit everyone, so we are talking to every fifth person on your block (in this neighborhood, on our list, and so on)," usually will satisfy the respondent. Sometimes assurance is needed that names will not be used in the report, or that the interviewer is not selling a product.

An alert interviewer usually can pick up something of common interest to talk about if a person shows signs of wanting to cut off the conversation. If a homemaker says she is busy, you can recognize the fact, express understanding, and make an appointment for a return call at a more convenient time.

Refusals should not be accepted except as a last resort. If an interviewer has a high number of refusals, he may need some help with his techniques. On the other hand, it might be in the best interests of the study to relieve him of his interviewing assignments. Sometimes, the person whose sensitive nature makes him a poor interviewer can make a valuable contribution to the study as a discriminating checker or, later, as a clerical aid in tabulating the data.

Almost anyone can be interviewed if you use the right approach. Do not force the interview on a person, and never become impatient or angry. If he refuses, try to make an appointment for some other time. If this also fails, thank him for considering your request and leave, showing no resentment. This leaves the way clear for another interviewer—the crew leader, an older (or younger) interviewer, or one of the opposite sex—to make a try at getting the interview.

Quite often, the respondent will give an interview when another interviewer approaches him—possibly because he now realizes the importance of the study or, maybe, because he regretted his original refusal but didn't want to back down.

Refusals happen very seldom so do not anticipate them. Try to avoid them but, if you can't, accept them graciously. Then forget it. Remember, there will be some refusals in any survey, so don't take it personally if you get one.

Before giving up on assignments, make at least two call-backs where the respondent was not at home on your first call, or was too busy to be interviewed. Be sure to keep a record of incomplete interviews, explaining what you did, and why you had to give up on them.

Securing Information — Thorough familiarity with questions helps to make a smooth interview. Under no circumstances may interviewers change the wording of questions. However, you can reread them and define or explain a word if you feel the respondent does not understand it.

Explain that you are going to write down replies, and why. If you read back answers to free response questions you assure respondents that they are not being misquoted. This sometimes helps to secure more complete replies as respondents may add to their original remarks if they desire.

Do not be afraid to wait for a reply. Time may be needed to form thoughts and to think of ways to express them. A smile, or "That's all right. Take your time in answering," sometimes helps. Be a good listener. Let the respondent talk. Do not rush him but try to discourage talk about incidentals if they have no relation to the study.

Be careful to follow questions in the order of their appearance on the schedule and never omit any that apply to the person being interviewed. However, be alert to notice when a person answers a question during conversation that is to be asked later.

Replies must be encouraged without appearing to agree or disagree. Agreement may influence a person to give future answers in terms of what he or she thinks you want. Disagreement may cause a person to suppress real opinions. Watch for reactions from the respondents; make marginal notes about them.

If interest appears to lag, some statement as to the time required to complete the interview may help. A change in tone of voice, or a shift in pace may help. Transition statements from one area of a schedule to another also help to maintain interest. For example, if you have been asking questions about clothing, you might say, "You have answered the questions on clothing. Now I have a few others about foods." Or, "The information you have given us about crops raised last year will be very helpful. Our next questions are about livestock."

In opinion or attitude questions, it is very important to assure respondents that there are no right or wrong answers. Frequently a person will ask an interviewer, "Is that what you wanted me to say?" In this case, you can say, "Yes, if that is your opinion. There is no one answer to this question."

More skill and patience are required to obtain replies to free response (open-end) questions than to checkoff questions. A friendly tone of voice and a conversational manner will encourage a reply. "I don't know," or "I have no opinion," are easy ways to avoid expressing real feelings. An interviewer must find ways of getting past such replies to get usable answers. Sometimes repeating the question in a different tone of voice helps. Sometimes careful probing is justified.

If probing is used, you must be careful not to use suggestive probes. For example, you might ask the question, "Why do you like a 15-minute radio program on one subject?" A suggestive probe would be, "More interesting?" This suggests a reply, and may be pounced upon by the respondent regardless of what his real reason might have been.

Once the person has started to talk, the interviewer may probe for more complete answers by comments such as, "Anything else?" "Other ideas?" "Another comment?" or "Are you sure that is all?" These do not suggest answers but do encourage the respondent to come up with a more complete answer.

It should be noted here that there are occasions when "no opinion" and "don't know" replies are valid answers. If a person really has no opinion or does not know, he should not be forced into replying further. More harm than good may be done by trying to force an answer if the respondent is reluctant to reply to questions. Rather than antagonize the respondent, the interviewer simply writes "refusal" or "did not choose to answer" beside the question.

A final reminder while securing information: Write clearly and make all entries as complete as possible. Although they may seem perfectly obvious when recorded, sketchy notes, unfinished phrases, or single words often prove to be meaningless or ambiguous to the person who must tabulate the answers. While reading back the reply to the respondent as a check on accuracy and completeness, try to imagine what a person, who must react to the recorded entry only, will get out of it.

Closing the Interview—A pleasant impression of both study and study personnel should be left with the respondent. He or she should be left with a feeling of having been helpful and that the cooperation has been appreciated. Always extend some expression of appreciation of the close of the interview.

Try to have the parting be such that your return, either as an individual or as a member of this or another research team, would be welcomed. You never know when you may have to call back about some item, or follow through on some afterthought related to the study. It is a matter of good common sense and professional ethics to try to preserve friendly relations for future studies, whether by you or by others, which may involve your respondent.

Editing—Too much emphasis cannot be given to the importance of checking the completeness, accuracy, legibility and consistency of each schedule before you turn it in. Careless editing by the interviewer can bias the results of the entire study. Never fail to go over the schedule, either before you leave the respondent or before you move on for the next interview. In this way you can clear up discrepancies and correct them.

Be sure all identifying information is entered. Write out abbreviations, complete fragmentary sentences, rewrite hard to read passages, and expand sketchy remarks. Make marginal comments to explain apparently inconsistent or unusual entries. Sign the form so that others will know who to contact for further clarification or follow-up tasks.

Although your work will be edited again, both by you and by others, the editing you do at this point is likely to be the most effective in improving the overall quality of the study.

Interviewing Without a Schedule (questionnaire or record form)—Most extension studies are made with the aid of some kind of record form. When interviewers are involved, some kind of schedule almost always is used. However, there are situations where it is not convenient or advisable to use a standardized record form. Oftentimes, in exploratory research particularly, a carefully structured questionnaire or schedule will hinder more than it will help. This type of interviewing will not be discussed to any extent in this chapter.

"Unstructured," "focused," "depth," and "nondirective" are terms applied to interviews made without apparent aid of a record form. The interviewer, sometimes called the "operator" in this type of interviewing as he may work with a tape recorder, dictation machine or similar device, appears to carry on aimless conversation with the respondent. He or an assistant records everything that is said and analyzes the comments after they have been typed.

Actually, the interviewer-operator directs the conversations by carefully inserted comments of his own so that all the areas he desires to have mentioned or discussed get worked into the conversations. He may work with a memorized que sheet or he may have a topic outline before him while talking to the respondent.

Although more involved, and difficult to transcribe, these same techniques may be used for group interviews. No doubt, we have all seen them demonstrated in radio or television interviews when at least the announcer knows what information he wants to come out during the interview.

Although these types of interviews have their place, and are increasingly important in extension research, they have to be ruled out for much of our research because more skill and training is required of the interviewers. Special probing techniques must be mastered to get results without appearing to direct the interview.

Data are more difficult to analyze in this form from both a time and a skill-of-operator standpoint. They do not contribute so readily to statistical analyses and, as a result, are limited to preliminary or exploratory research rather than research to serve as a basis for generalizing.

We are becoming increasingly aware of a need in extension to go beyond the usual descriptive type of study we do so often, and work out ways of really getting at what people think, what they want, why they change, and how they are motivated. These "nondirective" techniques may be ones we will have to adapt to our needs if we would keep up with the demands of our work. It is hoped that more extension researchers will desire to become better trained in these techniques. At the moment they are relatively new, but there is some good literature available about them for those who are interested. If an extension worker is interested in the type of interviewing and doesn't have the needed training, other help should be sought.

Summary

For extension research purposes, we think of interviewing as the more personalized ways of securing information, the face-to-face or person-to-person types of data gathering. We also think of the kind of interviewing in which a schedule or questionnaire is used rather than when the more elaborate and demanding nondirective techniques are employed.

We believe it is important for a person directing research to be aware of what interviewer responsibilities are and we discussed some of these responsibilities in part. Next we mention some things to consider in training interviewers and include a suggested outline for a six-hour training program for interviewers.

Then we elaborate on the four steps of the interviewing task—making the contact, collecting the information, closing the interview, and editing the interview form. We close the chapter with a brief discussion of interviewer techniques for situations where routine interviewing methods are not effective. We suggest some of the reasons why the newer techniques, for the present at least, are not readily adaptable to extension research and evaluation, but we go on to state why this is the direction in which extension researchers must explore to find ways of dealing with the growing complexity of extension problems.

Chapter XIII

DEVICES FOR COLLECTING DATA

Gladys Gallup

A good deal of behavior can be observed through the senses. We recognize improvement in terms of accepted standards in a 4-H Club boy or girl or an adult leader by watching the person preside or by hearing him talk at a meeting at different times.

We can drive through the country and observe the increase in compliance with soil enrichment practices, the use of suggested varieties of grain, the improvement in livestock, and the betterment in farm and home buildings and yards. We look for evidence of favorable change in light of what we have been teaching. Even changes in intangibles such as attitudes, interests, and personality traits are expressed in overt behavior. Such observation is good, but it is not sufficient.

The Need for Measuring Devices

Different kinds of devices have been designed to measure individual behavior. These are devices which for the most part gather evidence as to a person's knowledge, attitudes, or his activities, put in another way, what a person knows, feels, or does. These devices take but a few minutes to apply, whereas direct observation of behavior by competent observers often requires many hours or days and is influenced by observer experiences and understandings.

The questionnaire or other device used to evaluate a program should be worked out especially for that particular program. We cannot usually expect to find one ready made because differences in objectives or in the methods of approach generally make a technique that is effective for one program unsatisfactory for another. However, before constructing a new device, it is helpful to examine some of those which have been used previously. They may be a source of suggestions for the particular problem upon which we are working and may furnish test patterns for construction.

The Selection of the Measuring Device

When selecting the kind of devices to be used, we should keep two things in mind: (1) The teaching objective, and (2) the evidence for which we need to look in order to learn what progress has been made in reaching the objective.

In the educational field there are numerous devices that can be used to measure progress toward an educational objective. Some of those which are designed to observe and record behavior are:

Knowledge Questions—It is important to know whether a person knows information and the level of information.

Understanding Questions — It is important to know whether a person understands or can apply certain acquired knowledge of principles.

Skill or Performance Ratings—These are used to determine skill being performed or the product resulting. They may show the amount of skill the homemaker has or to what degree skills have been acquired.

Attitude Scales—These are used to show how people feel toward things, whether they are for or against certain questions, social issues, or problems.

Interest Checks—These are used to find out in what people are interested.

Confidence Testing—A measure of self-assurance, the feeling that you can do a certain thing.

Adoption of Practices—The adoption of a recommended practice is frequently used in extension research as evidence that the farmer or homemaker has changed behavior and acquired new knowledge, skill, and understanding.

Some Illustrations of Devices

Knowledge Questions or Checks on Levels of Information—Obviously it is essential to impart many facts in order to teach any subject matter in Extension. When we test we can ascertain rather easily whether the persons taught acquired this knowledge. But can be assume that they understand those facts which they have learned? Certainly we should not assume that they are able to apply their knowledge. If our objective is to develop understanding it is necessary to extend our questions in order to learn if this knowledge is comprehended and applied.

In the schedule used in a Wisconsin Study[2] are some illustrations of questions which attempt to get at knowledge only.

1. Some fertilizers are higher in plant food than others. Which one of these fertilizers has the most nitrogen? (Check one only)
   ( ) a. Anhydrous ammonia
   ( ) b. Ammonium nitrate
   ( ) c. 0-20-20
   ( ) d. 10-10-10
2. The acidity or sourness of a soil is sometimes expressed as a pH number. Which soil is the most acid? (Check one only)
   ( ) a. pH5
   ( ) b. pH9
3. Acid soils have a number of bad effects. Which one of these things does an acid soil do? (Check one only)
   ( ) a. Ties up potash so crop can't get it
   ( ) b. Ties up phosphorous so crop can't get it
   ( ) c. Makes soil hard
   ( ) d. Keeps soil from holding water
4. How many pounds of nitrogen would there be in 100 pounds of 3-12-12 fertilizer? (Check one only)
   ( ) a. None
   ( ) b. 3 pounds
   ( ) c. 12 pounds
   ( ) d. 27 pounds
5. Some crops have "hunger signs" which show that they are in need of plant food. When alfalfa leaves have white dots around the margins and the lower leaves dry up, what plant food are they probably short on? (Check one only)
   ( ) a. Nitrogen
   ( ) b. Phosphorous
   ( ) c. Potash
   ( ) d. Calcium
6. You've probably heard or read that you shouldn't plant corn on a 9% slope. How steep is 9%? (Check one only)
   ( ) a. 1 foot drop in 9 feet
   ( ) b. 9 feet drop in 1 rod
   ( ) c. 9 feet drop in 100 feet
   ( ) d. None of these

An example of a knowledge question in the field of marketing information for consumers was taken from the Lake Charles, Louisiana Study.[1] This question was designed to get at the level of information of homemakers in Lake Charles.

7. If you were going to buy a higher grade cut of beef, what is the main thing you would look for? (Check one only)
   ( ) a. Leanness, or free from fat
   ( ) b. Presence of a yellow rim of fat
   ( ) c. Deep red in color
   ( ) d. Little lines of fat running through the meat
   ( ) e. Don't know

This question was asked as an open end question. The enumerator then checked the corresponding answer on the schedule.

Other examples of knowledge questions are taken from the Arkansas Extension Pilot Cotton Project[11] on disease control.

8. Angular leaf spot is carried primarily from season to season by which of these means? (Check one or more that apply)
   ( ) a. On the seed
   ( ) b. In the soil
   ( ) c. Plant refuse
   ( ) d. Mechanical equipment
   ( ) e. Don't know

Below is a series of questions about common cotton insects.[11] The interviewer gave the respondent a picture of a common cotton insect and asked questions about it. The interviewer then checked or wrote in the replies on his record form.

9. What is the name of this insect?
   ( ) a. Bollworm
   b. Other ______________________
   ( ) c. Don't know
10. At what stage of growth of the cotton plant is it recommended that farmers start spraying or dusting to control this insect?
   ( ) a. Pre-square
   b. Other ______________________
   ( ) c. Don't know
11. What is the standard or indicator for determining when control measures for this insect should be applied?
   ( ) a. 10 to 25% punctured square — dry year
   ( ) b. 25 to 40% punctured square—normal year
   c. Other ______________________
   ( ) d. Don't know
12. What kind of insecticide is effective in controlling this insect?
   a. ______________________
   ( ) b. Don't know
13. How often should this insecticide be applied to effect continued control of this insect?
   a. ______________________
   ( ) b. Don't know

14. How much of this insecticide should be applied per acre, midseason; how much late season?

	Midseason	Late Season
Dust	a. ______	d. ______
Spray	b. ______	e. ______
Don't know	c. ______	f. ______

15. When do you quit applying insecticide for controlling this insect?

a. ______________________

( ) b. Don't know

Understanding Questions—It is important to find out whether a person understands or can apply certain acquired knowledge.

Comprehension or understanding usually can be inferred from what the respondent says or does. Present him with a new situation and see whether he can apply the basic principles taught him. The situation must be natural and not so new but that he can see the similarity to former situations.

There are various levels of understanding and degrees of understanding. Trying to measure at too high a level or too ideal a level will be disappointing. On the other hand, to measure at too simple or too low a level will not tell you whether the person can apply the knowledge at his own level of living and working.

Here are some illustrations of methods used to test understanding:

1. Teaching selection of accessories for the basic dress or suit.[7] The home demonstration agent or a local leader wears a basic dress or a basic suit and teaches the selection of accessories to be worn with it.

   After the teaching is finished a second assortment of accessories, including hats, shoes, scarfs, jewelry, purse, and the like, are brought before the group. The members in attendance are asked to select appropriate accessories for the agent or leader to wear for two or three different occasions. This will indicate if they know the principles and if they can apply them.

2. In teaching farmstead planning[8] in workshops to those farmers and homemakers interested in building, there are certain principles taught as to location of buildings on the farmstead.

   The extension worker asks each husband and wife present to draw a rough plan of their farmstead, showing slope of ground, direction of wind, view, location of highway and driveway. Then, using circles to indicate buildings, they put in the house and the farm buildings, applying the principles of farmstead planning. This tests their knowledge of principles and the ability of the husband and wife to apply principles to their own situation.

3. In teaching art principles[9] and the application of art principles in design and color in the home at the local leader training meeting, a mimeographed sketch of a living room including doors and windows is given each leader. Each person at the meeting is asked to arrange the furniture according to principles taught. Then swatches of materials and wallpaper are used to work out the color scheme for the room.

4. Problem-solving devices[10] can be used to measure understanding. The following is an example of the type of questions that can be used by 4-H Club boys and girls:

   A farmer wishes to market his spring crop of hogs before the usual drop in market price. Which of the following practices will help?

   (Check one or more)

   ( ) 1. Have pigs farrowed in March or early April.
   ( ) 2. Have pigs farrowed late in May.
   ( ) 3. Put brood sows and litters on pasture with self-feeders of shelled corn and tankage in separate compartments.
   ( ) 4. Wean the pigs at 6 weeks of age.
   ( ) 5. Provide self-feeders after weaning.
   ( ) 6. After weaning, put pigs in dry lot and feed entirely on corn.
   ( ) 7. Limit the feed during the summer and have hogs on fall crops of corn.

   Reasons: (Check the statements below which represent your reasons for choosing the practices you checked in the preceding list.)

   ( ) a. The usual market drop which affects the spring crop of hogs starts in September.
   ( ) b. The usual market drop which affects the spring crop of hogs starts in November.
   ( ) c. Under favorable conditions hogs will be ready for market at 6 months of age or less.
   ( ) d. Hogs gain most rapidly when hand fed.
   ( ) e. Hogs can usually be produced most cheaply on a limited grain ration for the first few months, followed by heavy feeding.
   ( ) f. Brood sows and their pigs are able to balance their own ration when provided "free choice" of tankage and corn.

Skill or Performance Ratings—There are two reasons for measuring skills: (1) To find out if the people have acquired a particular skill that has been taught and (2) to find out how many skills have been attained.

Skills that are important for the clientele to learn are listed, criteria are set for measuring the skills, and some devices are prepared for use.

Usually score cards and rating scales are used to measure skill, performance, and ability. They are used extensively in Extension at fairs and other events in judging products and in setting standards of quality. The product or skill to be rated is carefully analyzed, each aspect is considered separately.

It is necessary not to have too many points on a score card or rating device and there should be a description of two or more levels of quality for each aspect of the product, skill, or ability to be rated.

The examples below are devices for judging selected skills in a 4-H Club clothing project. With each skill is explained exactly what the "rater" should look for in rating that particular skill.

1. Score Card for Judging Garments

(Adapted from device constructed by Lois Corbett, Michigan Extension Service, for judging 4-H Club clothing projects.)

Club Member________________ Age_____
Local Leader____________ Rater____________

Directions: Rate garment on each item which applies to it, rating 1, 3, or 2, depending upon whether the quality corresponds to the description in the left hand column, or the right hand column, or falls between the two. Add scores at bottom of page.

Two or three items are shown for "construction of a dress."

STATEMENT No.		1	2	3	SCORE
20.	Skill in machine stitching	Crooked, too far from edge; not alike on both sides; stitch improper length		stitching even, very close to edge, alike on both sides, stitch proper length	

2. Judging Clothing Skills

These are the instructions to the local 4-H Club leader: "This chart is prepared for use in judging the sewing skills the girls are developing in their 4-H clothing work.

"The sewing skills of each girl are to be judged twice, once at the beginning of the project (or as soon as you have an opportunity to observe her skill and make a judgment) and a second time toward the end of the project.

"Make your best judgment of each skill for each girl using the following yardstick:

Yardstick for Judging Clothing Skills
(Five Degrees of Skill)

(1) **Practically no skill**—does not know how.
(2) **Little skill**—needs much help and practice.
(3) **Fair skill**—but still needs some help and practice.
(4) **Good skill**—needs very little help or practice.
(5) **Excellent skill**—does not need any help or practice.

"In the example below and on the following pages the numbers in the squares refer to the numbers on the yardstick. Place a circle around the number which shows your judgment of the girl's skill. The top row of the pair is for the beginning of the project and the lower row is for the end. The difference between the two ratings shows how much the girl has developed that skill.

Example: Skill in threading needles. (Consider how thread and needle are held and speed in threading needle.)	Beginning	[✓ 1]	2	3	4	5
	End	1	2	[✓ 3]	4	5

"This means that at the beginning of the project this girl had "practically no skill"—she didn't know how to thread a needle. But during the clothing work she improved until at the end she had developed "fair skill"—but still needed some practice and help.

"Some girls in some skills may have "fair skill" (3 on the yardstick) to begin with and go as high as "excellent skill" (5 on the yardstick) by the end. Others may make no improvement at all. Still others may begin at "practically no skill" and end at "excellent skill." Others may begin the work this year with "excellent skill" and of course end with "excellent skill."

"Be sure to base your judgment each time on your observations of the girl doing the skill and on the quality of work she does. Since some girls may not happen to perform some of the skills, you will not be able to judge every girl on every skill. Use a separate pamphlet chart for each girl."

3. Local Leadership Rating Scales

Extension workers have used different types of rating scales with local leaders. A point rating scale is sometimes used which indicates a scale of specific needs: 1 to 5 standing for very much, much, considerable, some, little or none, respectively.

For example: How much help does this leader need in conducting a discussion meeting?

1	2	3	4	5
Very much	Much	Considerable	Some	Little or none
Amount of Help Local Leader Needs				

Attitude Testing—The extension worker is concerned about the attitudes and beliefs of people. Sometimes people indicate a disposition for or against something, or sometimes they indicate uncertainty. A person may be favorably inclined toward high rigid price supports, he may favor flexible price supports, or he may be in doubt as to which he prefers. Attitudes and beliefs can be identified only when people are in a situation in which they are confident that they are free to expresss themselves as to what they feel or think.

"State Extension Workers' Ballots" illustrate an attitude device employing multiple choice questions. The extension worker can express how he feels on a number of questions. There may be several questions. The first question is more general, the last question more specific.

The ballot gives a person an opportunity to express his views on a number of questions related to the Extension Service. Directions read as follows:

"Below each question you will find five possible answers. Read all five answers, and choose the one that comes nearest to expressing the way you feel about the question. Then vote by putting an "X" at the left of the answer you have chosen."

"Vote for only one answer to each question. If none of the five answers to a question happens to agree exactly with your own opinion, vote for the one that comes closest to saying what you think about it. If you want to write any comments on the ballot, that will be all right; **but be sure to vote on one question** by making an "X." Use the back of the pages for your comments if you need more room to write. You do not need to put your name on the ballot."

1. How do you feel about the Extension Service as an institution?
   ( ) a. There are many serious weaknesses in the Extension Service.
   ( ) b. The Extension Service has some weaknesses, but none are too serious.
   ( ) c. The Extension Service has about as many good points as bad ones.
   ( ) d. The Extension Service has more good points than bad points.
   ( ) e. The Extension Service has no serious weaknesses.

2. How much help do you get from your District Supervisor in doing your work?
   ( ) a. He never helps at all in my work.
   ( ) b. He hardly ever helps.
   ( ) c. Is sometimes fairly helpful.
   ( ) d. Is usually quite helpful.
   ( ) e. Always gives me all the help I could possibly want.

3. IN MY OPINION, separate 4-H Club subject-matter leaflets and bulletins need to be prepared for use in 4-H Clubs, rather than using materials prepared for adults:
   Yes (1)____
   Yes, depending upon age of members (2)____
   Yes, depending upon subject matter content (3)____
   No (4)____
   No opinion (5)____

4. Now we would like to have your opinion on some ideas. I shall read eight statements to you and would like for you to tell me whether you agree or disagree with the statements. You may not have an opinion on some of them.

AGREE	DISAGREE	NO OPINION OR UNDECIDED	STATEMENTS
			1. Hogs as well as cattle may be affected by Bang's Disease.
			2. A cow gets more food value from a pound of ground roughage than from a pound of unground roughage.
			3. DDT is effective for controlling hog lice.
			4. If calves are to be dehorned, it is best that they be dehorned before one week of age.
			5. Barnyard manure gains in its fertilizer value if allowed to rot before being spread.
			6. Tuberculosis may be transmitted from chickens to hogs.
			7. It is often necessary to add phosphate to a barnyard manure to make the manure a well balanced fertilizer.
			8. A well rooted stand of Birdsfoot Trefoil makes an excellent permanent pasture grass.

Interest Checks — An extension program may succeed or fail in proportion to the emphasis which the extension worker puts upon the development and encouragement of interests. Interests indicate likes and dislikes, satisfactions and dissatisfactions.

Interest checks, devices for finding out in what the people are interested, have been used to considerable extent in Extension.

The "Suggested Form for an Interest Questionnaire" illustrates a method of finding out what help rural people are interested in getting from the county agent. This is used as a partial basis for program planning.

1. **Form for an Interest Questionnaire**[6]

In the following list of items, the interviewer will check for each the degree to which the person interviewed is interested in receiving information or other assistance, and methods preferred.

Item	Degree of Interest Check ( )			Methods Preferred
	Much	Some	None	
1. Selection of a good dairy cow				
2. Organization for community improvement				
3. New varieties of seed				
4. Outlook information on crops and livestock				
5. Methods of pasture improvement (90 items were included in this list.)				

(All extension methods will be keyed as follows and each enumerated for person to select preference:

M—Meeting
V—Visit
C—Circular
B—Bulletin
N—News story
Etc.)

"How to Decide If You Are Interested in a Subject—

"**Much interest** means: Ask yourself whether you **would read about it right away** if you had the reading material, or if you **would attend a meeting with other farmers** to hear the subject discussed. If you would, **then check in the "Much" column,** meaning **much** interest.

"**Some interest** means: If you are not sure, or if you would probably put off for some time reading about it or attending a meeting to hear it discussed, then check in the **"Some" column,** meaning **some interest,** but **not much.**

"If you feel **sure you would not read about the subject** or attend a meeting on it, no matter how much time or reading material you have, or how many meetings were held, then check in the **"None"** column, meaning **no interest.**

"Be careful to consider and check every statement, but do not spend too much time on any one statement."

2. How long have you lived in this community?
a. ________ No. years

3. How satisfied are you with the way local government affairs are handled? Would you say you are not at all satisfied, not very well satisfied, pretty well satisfied, or very well satisfied?

( ) a. Very well satisfied
( ) b. Pretty well satisfied
( ) c. Not very well satisfied
( ) d. Not at all satisfied
( ) e. Undecided
Any comments: ______________________

The above question is one that has been used to appraise satisfactions.

Confidence Testing—Self-assurance or self-confidence, the feeling that one can do a certain thing has in the past been considered too intangible to measure. And yet you have probably observed a person and noted that "he is very confident" or "he lacks confidence." In so doing you have observed and drawn a conclusion; you have "measured" his confidence. Therefore, there must be tangible evidence of confidence.

The following is a device or question which has been proven valid and has worked successfully. It can be adapted to measure confidence in many different kinds of things.

Have you ever made a school dress for yourself?[7]

( ) a. Yes
( ) b. No

Do you think you can? (Check **one**)

( ) c. No, I don't think I can.
( ) d. Yes, I think I can, but am not sure.
( ) e. Yes, I feel sure I can.

Practices Adopted or Action Taken — Perhaps most of the devices and questions used in Extension to measure change are aimed at finding out if the person has adopted the practices taught. A practice adopted is also evidence of knowledge acquired and attitude changed. In time and energy management certain principles are taught such as "sitting to work" and "eliminating some jobs."

Following are examples of two questions taken from a check sheet prepared by a home management specialist. This check sheet was designed to be used before and after the learning experience.

"There are several different things which people do around the house. We are interested in finding out how many women are doing these things. So I would like you to go through the list and check which ones you do."
Home economics example:

1. In ironing, do you:

	Usually	Sometimes	Never
a. Iron towels?	(1)____	(2)____	(3)____
b. Iron sheets?	(1)____	(2)____	(3)____

2. Do you sit to:

a. Iron?	(1)____	(2)____	(3)____
b. Prepare vegetables?	(1)____	(2)____	(3)____
c. Wash dishes?	(1)____	(2)____	(3)____

Agricultural example:

3. Do you sell eggs on a graded basis?
   ( ) a. Always
   ( ) b. Sometimes
   ( ) c. Never

4. Where are eggs kept until marketing?
   ( ) a. In the house
   ( ) b. Where ____________________
   ( ) c. Outbuilding (specify) ____________

Practices Adopted and Methods Responsible — Many of the questions determine practices adopted and attempt to determine methods the respondent thinks are responsible.

1. Do you have electricity?
   Yes (1)____ No (2)____

2. If yes, a. Do you, yourself, repair your own electric cords or plugs?
   Yes (1)____ No (2)____
   b. If you repair own cords or plugs, where did you learn to do it? ____________________

Summary

Devices for collecting data are forms on which the data can be recorded and used at a later time for analysis and tabulating. Memory does not need to be depended upon.

In selecting a device we need to keep in mind the objectives to be evaluated and the evidence to look for. We have given some examples of questions and forms for eight kinds of objectives. Many times, however, we must construct our own devices. This is the subject of the next chapter.

References

1. Cooley, E.; Powe, D. D.; Matthews, J. L.; and Ueland, G. **Informing Consumers in Lake Charles: An Evaluation.** Baton Rouge, La. Agr. Ext. Pub. 1233, 1958.
2. **Factors Influencing Knowledge of Technical Soils Concepts by Wisconsin Farmers,** Bulletin 27, College of Agriculture, University of Wisconsin, March, 1957.
3. Fessenden, J. G. "Tools for Testing." **Extension Service Review,** September, 1955, pp. 178-179.
4. Fleck, H. **How to Evaluate Students.** Bloomington, Ill. McKnight and McKnight Publishing Co., 1953.
5. Frutchey, F. P., Deyoe, G. P., and Lathrop, F. W. "The Measurement of Understanding in Agriculture." **The Measurement of Understanding.** Chapter XIV, 45th Yearbook, Part I, National Society for the Study of Education, 1946.
6. Frutchey, F. P. and Church, H. L. **Educational Growth in the 4-H Clothing Project, Missouri,** USDA, Ext. Ser. Cir. 382, Washington, D. C., 1942.
7. Frutchey, F. P. and Forbes, M. E. **Evaluation in the 4-H Clothing Project, Massachusetts, 1940-41,** USDA, Ext. Ser. Cir. 384, Washington, D. C., 1942.
8. Leagans, J. P. **The Educational Interests of Farm Operators in North Carolina as Related to Work of the Agricultural Extension Service.** Ithaca, Cornell University, 1948.
9. Noyes, H.; Dickerson, L. B.; and Lusk, M. **Examples of Problems or Situation Statements Used in Evaluation.** Extension Service, Washington State College, Pullman, Washington, 1955.
10. Parten, M. B. **Surveys, Polls, and Samples.** New York, Harper and Bros., 1950.
11. Raudabaugh, J. N.; Saunders, J.; Cavender, J. **Arkansas Extension Pilot Cotton Project.** (In progress.)
12. Rothney, J. W. M. **What Research Says to the Teacher.** No. 7. "Evaluating and Reporting Pupil Progress." National Education Association, Washington, D. C., 1955.
13. Sabrosky, L. K. **Six Keys to Evaluating Extension Work,** USDA. Ext. Ser. Cir. PA-371, Washington, D. C., 1958.
14. Sabrosky, L. K.; Leagans, J. P.; Gallup, G. **Extension Evaluation, What Is It, How It Is Done, Its Use.** Outline for Summer School Course, including discussions and studies developed. Colorado A. and M. College, June 24—July 13, 1946, USDA, Federal Extension Service.

Material in this chapter is based on a chapter, "Selection of the Measuring Device," prepared by Laurel K. Sabrosky in Reference No. 14 above. Also the chapter is based on subsequent discussions with Mrs. Sabrosky and other members of the Division of Extension Research and Training.

Chapter XIV

CONSTRUCTION OF AN EVALUATION DEVICE

Gladys Gallup

The selection and construction of the measuring device depends upon the kind of evidence wanted in light of the objectives of the study. At this point the evaluator needs to think ahead how he plans to collect the information.

The county agricultural agent may decide to use an observation sheet to carry with him as he makes visits to a cross section of farms and homes. He may decide to send mail questionnaires to a sample of farmers or homemakers. The home demonstration agent may develop report forms such as checklists or questionnaires to use in collecting information regarding the adoption of practices from a sample of participating homemakers.

Sometimes a field study is made. This is a more formal evaluation study. It requires a different type of questionnaire or schedule as it is to be used to interview a cross section sample of the people living in the county. You need, therefore, to consider not only the kind of evidence which you want to meet the objectives of the study, but also how you plan to collect the information.

The questions asked should be so set up that information or data collected separately from a number of people can be added together or summarized to present a true picture of the group.

The questions in the measuring device should be definite and specific questions related to the major questions which are to be answered by the evaluation. For example, if one of the major questions to be answered by the evaluation is "What has been accomplished to control soil erosion on the farms," specific questions relating to soil erosion control would be constructed for the questionnaire.

When the replies to these specific questions are summarized, they answer the major questions on soil erosion control. The major questions are the main purposes of the evaluation and the specific questions of the questionnaire are designed in relation to the main purposes of the evaluation.

Criteria for Constructing a Measuring Device

For a device, such as a report form, an interest check, or a questionnaire to be effective there are five criteria which it must meet. These criteria are validity, reliability, objectivity, practicability, and simplicity.

Validity is concerned with measuring what you want to measure. A valid device actually gets evidence of progress in the teaching objectives or of the information you desire. The two elements of validity are behavior validity and subject-matter validity.

Behavior Validity—A device has behavior validity if the behavior called for in the questions is the same as that expressed in the teaching objectives. The device gets evidence of the behavior you are teaching. Measuring skill, for example, will require a different device from measuring knowledge, because of the different behavior involved.

Subject-Matter Validity—A valid device must deal with the particular subject matter of the teaching objective. It is obvious that if we are measuring attitudes toward soil erosion control, we would not ask questions about attitudes toward using fertilizer. The device must contain the subject matter and situations in which the behavior is expressed as called for in the teaching objectives.

When a device is valid from the standpoints of both behavior and subject matter, valid evidence can be obtained from which to draw conclusions and make recommendations about the teaching objectives, methods or other aspects of the program which we are trying to evaluate.

Reliability is the degree to which the sample of people or of subject matter represents and gives results which are like the results that would have been obtained from the whole population of people or of all the subject matter. It is a measure of whether the random sample is large enough to give sufficiently stable results. The sample may be a sample of people, a sample of subject matter, or a sample of behavior over a period of time, depending upon the "population" about which interpretations are to be made. Extension research has been concerned mainly with samples of people.

Objectivity means having a concrete, observable, and tangible basis for making a judgment or interpretation. It is in contrast to **subjectivity,** opinions, or guesses. An objective measuring device is so constructed that the results from it are not influenced by the difference in people who administer or interpret it. If equally competent people ask a question, interpret the answers, and get closely similar results the measuring device is considered objective. The use of numbers helps to make interpretations more objective. For example, "100 persons" is more objective than "many persons."

Practicability—A measuring device must be practical to use. It is necessary to consider time, cost, and convenience. Will the form take more time than it is worth; will it cost more than you can afford? It must be within your resources considering all the demands made upon you. In planning the time schedule, the length of the questionnaire, the time involved in summarizing it, the cooperation you will need of the people and the probable value of the results obtained continually must be balanced against its costs.

Simplicity—Is the report form or questionnaire simple, easy to understand, easy to answer, easy to administer, easy to tabulate, and easy to summarize? The more simple a questionnaire, the more practical it will be to use and the time between field work and the report will be minimized. Simplicity should not be confused with brevity.

The degree to which a measuring device meets the criteria is conditioned by the circumstances of the study and the use to be made of the findings.

Types of Questions

In wording questions for the different measuring devices used in Extension simple types of questions have been used.

Dichotomous Questions. Dichotomous means division into two parts or division by pairs.

These questions present opposite choices designed to get response such as yes or no, agree or disagree, approve or disapprove.

The chief advantage of dichotomous questions is simplicity. They take relatively little time to ask, and give a clear-cut answer which can be recorded and tabulated easily.

The following are typical dichotomous questions from extension studies:

Do you happen to know who is the county agricultural agent? Yes ____ No ____

Although these questions are designed as "dichotomus" they usually should and do allow for a third response, such as "no opinion," "not sure," or "don't know." For example:

During the past year have you attended:

A women's extension club meeting conducted by the county home agent?

Yes ____ No ____ Don't remember ____

The chief advantage of the dichotomous question is that it can be answered with a concrete, definite answer. It gives a clear-cut reply which can be easily recorded and tabulated.

One of the limitations of the "yes" or "no" question is that a slight misunderstanding of its meaning may result in a complete removal from a positive to a negative answer. This may also increase the "no answer" rate. This is especially true when getting a person's opinion.

Multiple Choice Questions are used extensively in extension questionnaires. A direct question is asked followed by a number of possible responses. The person is asked to select from several clearly stated alternatives.

Multiple choice questions are used when there can be several responses to a question. The person answering these questions is given an opportunity of choosing one or more of several possible answers that are true, or represent his opinions or ideas, or come closest to them.

The directions may indicate that only the correct response should be checked or the directions may indicate that the respondent can check as many responses as he believes apply.

Multiple choice questions are not easy to construct. The possible answers must be plausible and the right answer must not be overly conspicuous. Example of multiple choice questions where one answer is correct:

Which one of these plant foods put on as commercial fertilizer will "leach" or wash out of the soil most easily? (Check one only)

nitrogen ____
phosphorous ____
potash ____
calcium ____

An example of multiple choice checklist type in which one or more answers may be checked:

Which of the following methods do you use for keeping up with **your farm business?** (Check as many as you use)

a. ____ Keep bills in a drawer, envelope, or some convenient place.
b. ____ Keep check stubs.
c. ____ Keep a ledger (book) which I developed on my own.
d. ____ Keep a farm cash account book (printed form).
e. ____ Keep in mind my expenses and receipts and add them up at the end of year.
f. ____ Other: ____

Open End Questions—These questions are usually used in personal interview schedules. The respondent is free to answer as he likes. Open end questions give such a variety of responses that the job of classifying and analyzing is both time consuming and difficult.

Open end questions are often used in pretesting to get a line on possible answers for an eventual check list type of question.

When it is necessary to employ questions which give the respondent free latitude in his responses, open end questions are used. Example:

1. How does your milk production in October, November, and December compare with other months of the year—is it higher, lower, or about the same? (Circle one)

IF "HIGHER" OR "LOWER"

2. About how much (higher) (lower)?__________
(percent)

IF FALL PRODUCTION IS SAME OR HIGHER

3. Many farmers are not able to keep their production up during these months. How do you manage to do it?____________________
____
____

4. Where did you get the idea?____________________
____

IF FALL PRODUCTION IS LOW COMPARED WITH REST OF YEAR

5. Have you ever heard of the idea of making your milk production even throughout the year by having cows freshen at different times? Yes____ No____

6. If yes, what do you think of the idea?__________
____

Wording of Questions

The wording of questions affects the answers. Each question must be worded to give exactly the information desired. Go over each question and ask yourself these questions.

1. Is the question related directly to the purpose of the study?
2. Can there be any doubt about the definition of any word in the question? Will each respondent understand each word as you do?
3. Are the words in each question familiar to the person who is to answer?
4. Are questions simple, short, clear, and concise?
5. Have you avoided asking leading questions—questions worded in such a way as to suggest answers?
6. Have you avoided "double-barreled" questions, questions that have more than one idea?
7. Have you avoided double negatives?
8. Are definite time limits on questions specified, for example, "During the past 12 months?"
9. Has a check **yes** and **no** answer been used whenever advisable?
10. Is there an opportunity for the respondent or cooperator to indicate that the question does not apply to him or her?

Face or Descriptive Data Questions

Besides the questions relating to the objectives of the work being studied, additional information may be needed about the respondent which will help in interpreting the answers—such items as number in family or place of residence.

Certain "face data" questions which can be checked against census data have been used in extension studies. These include: Formal schooling, place of residence, age groups, total net annual family income, and certain items on the Sewell level of living scale such as running water in the home, complete bathroom, and other conveniences.

These face data questions provide information by which the study data can be classified. For example, we may want to know if the people who adopted certain practices live on a farm or not, if they are older or younger, or 40 years of age and over, or under 40 years of age, if their formal schooling was 8 years or less, or more than 8 years. (A much more detailed classification may be desired.) By classifying the information in this way we can find the relationship between different factors and changed practices. Examples of face data questions:

Please check the HIGHEST grade you completed in school:

Elementary school
a. 1st to 4th grade ____
b. 5th to 6th grade ____
c. 7th grade ____
d. 8th grade ____

High school
e. 1 to 3 years ____
f. 4 years ____

College
g. 1 to 3 years ____
h. 4 years ____
i. more ____

The questions of this type that are to be included in the questionnaire depend upon the purpose of the study and how the data are to be analyzed. Other examples of face data are:

Where is your home located? Find the one of the seven following answers which best describes where you live. (Check one only)

a. On a farm or ranch ____
b. In the country but not on a farm or ranch ____
c. On a farm or ranch part of the year and in town part of the year ____
d. In a town or village with less than 2,500 people ____
e. In a town or city with 2,500 to 10,000 people ____
f. In a town or city with more than 10,000 people ____
g. In a suburban area (not in city limits) of a city with 10,000 or more people ____

In which age group do you belong? Please find the age group below in which you belong. (Check one only)

a. 19 years or under ____
b. 20—24 years ____
c. 25—29 years ____
d. 30—39 years ____
e. 40—49 years ____
f. 50—59 years ____
g. 60 years or older ____

Please indicate which of the following is nearest your net income after farm or business expenses are deducted. (Check one only)

a. Less than $1,000 ____
b. $1,000 — $1,499 ____
c. $1,500 — $2,499 ____
d. $2,500 — $3,499 ____
e. $3,500 — $4,999 ____
f. $5,000 — $6,999 ____
g. $7,000 — $9,999 ____
h. $10,000 or over ____

Core Questions

There is a series of questions that has been developed and used in extension studies to determine the knowledge which people have of extension, level of living items, and face data questions.

There are many advantages in using these core questions as they have been pretested under different conditions and are well designed. Most important, they provide comparable data because the questions are identical in wording. This insures that the results of a study will be more meaningful because they can be compared with the findings of other related studies.

In addition to the face data questions, a series of questions has been developed to determine knowledge of and participation in extension. Examples of such questions which deal with 4-H Club work are:

Do you know something of 4-H Club work?
Yes____ No____
Do you have children in 4-H Club work?
Yes____ No____
Have you ever had children in 4-H Club work?
Yes____ No____
Have you yourself ever been a member or leader of 4-H Club? Yes____ No____

Another example of a core question as used in home economics nutrition studies is:

Do you yourself drink milk? About how much each day?

a. One glass ____
b. Two or more glasses ____
c. None ____

Sequence of Questions

Logical Order—Place the questions in the logical order so that a person reading them does not have to jump from subject to subject in order to answer them. When the data are to be obtained by the interview method, the questions should be grouped so that the conversation will lead logically from one question to the next.

If general and specific questions are used, the general questions should precede the specific ones; otherwise, the respondent is likely to answer the general questions in terms of the earlier specific ones. For example, a general question as "During the past year what changes have you made in your house, such as remodeling?" should precede the more specific question, "During the past year what changes have you made in the kitchen?" You do not want the respondent to think about kitchens only and not mention other more general improvements.

Psychological Order — You can have a logical order of questions but the opening question should also have human interest appeal. The opening questions should be easy to answer and of interest to the respondent.

A few simple questions at the beginning of the questionnaire will give the respondent confidence in himself. This is invaluable for carrying the interview through later and more complicated questioning. Questions of personal nature, like asking for the age or income level of the informant should be placed near the end of the questionnaire.

In designing mail questionnaires the arrangement of questions in order of interest is particularly important.

Mechanical Setup

Four Parts to a Questionnaire—Each well-planned device or questionnaire has four parts:

1. Identifying information such as name of questionnaire or study, county, date, and name of person filling out form. If the study is of the

more formal type in which schedules are used by personal interviewers, the schedule should also include number of survey area, name of interviewer, and place for checker's initials. Also the schedule should include final disposition such as: O.K., refused, incomplete but closed, etc.

2. The **introduction** is important because it tends to orient the respondent to the purpose of the survey and tries to motivate answers to the questions. When records are gotten by mail, motivation is especially important. The opening sentences should explain:
   a. By whom the study is being made.
   b. What is the purpose of the study?
   c. To what use the information will be put.

Adequate response depends largely on how well-stated are the facts of the introduction. Any respondent is due the courtesy of an explanation covering the above points.

3. The content questions based on objectives.
4. Face data, if face data are necessary.

Physical Arrangement—The following six suggestions are on physical arrangement:

1. Number each item consecutively.
2. Never repeat a number on the schedule. This facilitates reference to items where instructions and definitions are given.
3. Have each item clearly distinguished for tabulation purposes.
4. Vary the type within a question to emphasize the most important words or phrases.
5. Make form easy to read and fill out. Have places for checks rather than for words even as short as yes and no.
6. Place spaces for such checks in vertical columns along the edge of the sheet to make summarization easier.

Adequate Space—When you expect information to be written in, provide adequate space for it. Do not expect people to use margins.

Good Appearance—The form should have good appearance. This can be accomplished by good paper, large print, wide spaces, and good arrangement.

Special Rules for Constructing the Mail Questionnaire

1. Appearance. The questionnaire should be attractive. This can be accomplished by appropriate use of colors, pictures, spacing, type, arrangement, and so forth.
2. Length. The questionnaire should appear easy to fill out and be as short as possible.
3. The letter of transmittal or cover letter sent with the questionnaire should have a personal touch and an appeal which seems likely to bring the highest proportion of returns as indicated in the pretest.
4. The letter should be frankable. With each cover letter and questionnaire there should be an enclosed franked envelope addressed to the extension worker for the respondent to return the questionnaire.

Pretesting the Questionnaire

Every device used in measuring should be pretested. The way to learn whether the test is ready for use is to try it out. The pretest helps you to know if the content of the questions and the form of the questions are satisfactory. There are two important points:

1. Try the questionnaire with persons who resemble those with whom you are finally going to use it.
2. Test the questionnaire under conditions similar to those in the study. For example, when pretesting a mail questionnaire have persons answer it without help from you. Tell them you have a questionnaire you wish them to try out and when they have finished ask for opinions and suggestions.

From a careful pretesting you get information about the following: How long it takes to fill out the questionnaire; whether the questions follow each other in natural order; whether ample space is provided for answers; and whether the directions are understood.

You have to determine how widely you pretest the questionnaire, and how often you revise it, in terms of importance of the study, the precision you need, and the amount of time and effort you can afford to spend upon the questionnaire.

Summary

The questions in the evaluation device should be definite, specific and related to the major questions to be answered by the evaluation. For a device such as a questionnaire to be effective there are five criteria which it must meet. These criteria are validity, reliability, objectivity, practicability, and simplicity.

The wording of questions affects the answers. Each question must be worded to give exactly the information desired. Special rules for constructing questionnaires should be followed.

References

1. Good, C. V., and Scates, D. E. **Methods of Research.** New York, D. Appleton-Century Co., 1954.
2. Hawkes, H., Lindquist, E. J., and Mann, C. R. **The Construction of Achievement Examinations.** New York, Houghton, Mifflin Co., 1936, Chapters 2 and 3.
3. Jahoda, M.; Deutsch, M.; and Cook, S. W. **Research Methods in Social Relations, Part Two: Selected Techniques.** New York, The Dryden Press, 1953, Chapter 12.
4. Parten, M. B. **Surveys, Polls and Samples.** New York, Harper and Bros., 1950.
5. Payne, S. L. **The Art of Asking Questions.** Princeton, N. J., Princeton University Press, 1951.
6. Sabrosky, L. K. **Six Keys to Evaluating Extension Work,** USDA, Ext. Ser. Cir. PA-371, November 1958. U. S. Government Printing Office, Washington 25, D. C.

Material in this chapter is based on a chapter, "Construction of the Measuring Device," prepared by Laurel K. Sabrosky in "Extension Evaluation and Course Outlines of Materials Prepared for Summer School Teaching of Extension Workers." Also the chapter is based on subsequent discussions with her and other members of the Division of Extension Research and Training.

Chapter XV

TABULATION

Ward F. Porter
Mrs. Jewell G. Fessenden

Tabulation is the process whereby study data are organized or grouped according to plan so as to facilitate analysis and interpretation. Most of us use the simpler tabulating or tallying techniques quite often in our day-to-day extension tasks. We note the number of incoming and outgoing telephone calls that are made at the office. We keep a running total on how many publications we distribute and what kinds. We have a tally sheet for office calls, farm visits, meetings with extension groups, liaison actions with other organizations, and the like. We have a frequency distribution of the topics we deal with on our radio or television programs or in our newspaper or magazine columns.

When it comes to evaluation work and research in general, it is obvious we could not get the job done without resorting to tabulation. We might not have to work out elaborate tally sheets or involved coding plans, but we must do some counting or we can make no pretense of having done an objective study.

Planning for Tabulation

We should work out our tabulation plans prior to the completion of the questionnaires or other record forms. We have already pointed out that we need to know what information we want before we can build a schedule or questionnaire to secure it. Thinking about the tabulation should properly commence after we decide what information we want and while we are working out practical and objective ways to collect the necessary data.

In planning for tabulation, we should keep four major considerations in mind.

1. **What Resources and Skills Will Be Available to Us?** We cannot plan for a very comprehensive tabular analysis if we must have a quick answer or if there will be limited clerical and secretarial help. If we are involved in the kind of study wherein we have to make our own analyses at odd moments stolen from our regular tasks, we have to be satisfied with relatively brief and simple study designs, record forms, and easily tabulated answers.

Where machine facilities and operators are available, we can plan much more involved and elaborate tabulations and can work out extensive cross-tabulations. In such cases, it is essential to include the machine technicians in our planning. Since many schedules are "pre-coded," the assistance of machine operators may be required well in advance of actual data collection.

If we will have to do hand sorting and counting with or without assistance, we cannot hope to take on so large a program of analysis as we would if we could count on electronic equipment. The problem that could be worked up in a few days with a trained crew of machine operators might take months, or be impossible to accomplish, using hand methods, before the findings would be obsolete.

In addition to physical facilities, we have to consider certain human factors, such as our own abilities and those of persons who will be at our disposal for the study. If we have no competence in statistics and have difficulty in following through on involved relationships, we may have to settle for reporting totals, working out a few percentages, and possibly calculating some averages. If we have some competence in such matters, or someone is available to guide us, we can work out a more comprehensive plan of analysis-ranking, correlating, intercorrelating, testing for significance, validity, reliability, and so on.

2. **What Information Will We Need for Our Analyses?** The purpose of the schedule or questionnaire, of course, is to secure the information we need. However, we can be more certain it will secure the information we need by drawing up some tentative tally sheets and work tables, preferably before the final draft of the schedule. This will help us see, for example, whether or not the probable answers will be in a form or of a type that will facilitate the analyses we wish to conduct. Once we are satisfied in this regard, we can make up the final tabulation forms, and be ready to make entries as soon as the data are received.

We must remember that we cannot plan for an analysis that requires more detailed information than can be supplied by our record forms. For instance, we cannot make year-by-year breakdowns in our tables if our age data are collected in terms of five-year intervals. Nor can we readily report average or mean income if we collect income data in terms of categories such as under $5,000, $5,000 through $9,999, and $10,000 or over.

On the other hand, however, we can make our tabulated information less detailed than that reported in our source. We may collect age information by age at last birthday and report it in our tally sheets by five-year intervals. We may have a complete age-sex breakdown but report in terms of numbers of adults or children per household. We may secure actual income data (very difficult to do) but report in a frequency distribution by $5,000 intervals.

3. **What Relationships Do We Wish to Study?** We must decide on what relationships we plan to study so that we have the proper questions in our data-gathering instrument and have adequate provisions in our tabulating forms to record the relationships in categories suitable to our needs.

To illustrate: If we want to know the differences in interests of respondents by age, educational level, place of residence, and sex, we must be able to sort on these items and then tabulate in terms of their answers regarding interests. In our tabulation plans we must decide what age, residence, and education categories will serve our purposes, and then be sure our source reports the information in enough detail to set up these categories.

4. **What Comparisons Do We Expect to Make?** Whenever we intend to compare certain data from our study with that of other studies, we must be certain to collect the necessary information and assemble it in such a way that there is comparability. Quite frequently we like to check on the probable representativeness of our sample by making comparisons with census or other statistical information related to our study area. Whenever this is our intention, we must be sure to have our background information (our "face data") defined and classified so as to match that used by the census, or other source to be used in our comparisons.

This is a strong argument for collecting our basic information in quite detailed fashion. Then we are at liberty to make a wider variety of tabulation decisions. We can group our data in the conventional categories used by the census and similar governmental agencies to make comparisons with their findings. Then, if needed, we can make special groupings to permit comparisons with findings of special studies by colleagues in Extension or elsewhere who have done work similar to our own.

Another common occasion for comparisons is in before-after or trend studies. When you are studying the change that has occurred since some previous reporting date, or when you are observing the effects of an action program, you have urgent need for comparable data. Whenever this is your objective, you must ask questions and prepare tabulations in essentially the same form for both occasions.

You cannot ask for annual family income one time and annual farm income the next. Nor can you ask for income to the nearest $1,000 one time and then decide to simplify the item next time by setting up $5,000 income categories.

Tabulation Techniques

Tabulation consists of listing or grouping answers from all respondents, or from all members of particular groups within the population being studied, so that summary answers can be obtained. Possible techniques for assembling these summary, average, or consensus answers vary widely. The ones you select will depend on the kind of study you are doing, the resources that are available, and what you want to get out of your data. We will mention some of the techniques that you might use in extension evaluation.

Tally Sheet—This is simply a counting aid, although it can be elaborated into a sorting device. On a card or sheet of paper of sufficient size, you set up column headings according to the categories into which you wish to sort or subdivide your data. Down the left side of the page you list the possible answers to the question.

County ______________________
Person(s) tallying ______________
____
Date ______________________

HAND TALLY SHEET (Example)

______________ Community Development Survey

	All household	Total farm	Total nonfarm
7. Level of living:			
Total households	44	25	19
No response	3	1	2
Reporting	41	24	17
a. Electricity	40	(23) 卌 卌 卌 卌 |||	(17) 卌 卌 卌 ||
b. Telephone	26	(12) 卌 卌 ||	(14) 卌 卌 ||||
c. Running water	25	(10) 卌 卌	(15) 卌 卌 卌
d. Central heating	19	(6) 卌 |	(13) 卌 卌 |||
e. Electric washer	27	(13) 卌 卌 |||	(14) 卌 卌 ||||
f. None	0	(0)	(0)

In our example, the sorting factor is residence (farm or nonfarm), while the question concerns the level of living of the families in the sample.

We will assume the respondents (or the interviewers) were asked to check only the facilities which were available in the homes.

Since it is necessary to account for each and every respondent, allowances have been made for nonresponses, and for households reporting none of the items listed. This "rule" applies in general to all tabulations. For certain types of questions, there might be justification for still another category—that is, one to account for respondents who report not knowing or "no opinion" or "don't know." The importance of this consideration has already been stressed in our discussion of questionnaire construction.

To construct such a tally sheet, you reserve space at the top for identifying information. Depending on the circumstances, such identifying data might include: (1) working title of the study, (2) working title of the particular table (tally sheet), (3) the name or initials of each person working on the sheet, including counter, reader, checker, and any others, and (4) date work was completed. It is a good idea to type the column and row headings and draw in necessary guide lines.

You can make the tally with a typewriter, if you have a wide-carriage machine, by having one person read off the numbers and where they should go, and the other enter them with the typewriter. The more conventional way is to draw up the tally form and then make the entries by hand with pen or pencil. You can still make typewritten copy for your permanent record, however.

In tallying, it is customary to make four perpendicular marks and then make the fifth one a slash through the previous four. This facilitates counting by fives and makes it easier to observe relationships by inspection rather than by actually counting. Subtotals can be placed inside brackets at the end of the count or immediately above or below.

This technique is most effective when only a few records are to be tallied, and when there is not a great amount of subsorting. The greatest disadvantages of this technique are the likelihood of error and the difficulty in checking for error. About the only way to check the work is to do it over, or have someone else do it over and compare results. Even the same results do not rule out the possibility that counteracting errors have been made. The prospects for error increase materially if there are interruptions during the tallying process.

Although tally sheets are most useful in limited, small-scale studies, there are at least two occasions when tally sheets have a place in the more complex evaluations or surveys. First, they can be used to advantage in taking off key items of interest in a preliminary, tentative, unchecked sort of way. Such data should never be reported or published, but they can be used to get an idea of what the results will be.

A second occasion where they can be useful is when the researchers have only vague ideas of how certain data should be classified or categorized. This occurs often with essay type items. The clerks can take a portion of the records, note each reply as it occurs, and mark down a tally for each time it is repeated. Soon a pattern emerges and you can set up sufficient broad categories (column headings) to take care of most of the answers you are likely to get. Any others can be scored in the "other" column. At this point, the researcher is in a position to work up a code for use in some one of the other tabulating techniques we will mention shortly.

Handsorting — This is another counting device which works quite well with small amounts of data. Instead of noting the answer and entering it on your tally sheet, you sort the answers into agreed-upon categories by stacking up record forms into piles containing the same or sufficiently similar answers. Then you count the forms in each pile to get totals for your record sheets, frequency tables, or charts. An alternate approach, where the record forms are lengthy or bulky, is to transfer the questions and answers to special answer cards. Both questions and answers can be coded or abbreviated to save space.

The major advantage of the method is its simplicity. You need no special equipment to do the work; and no special skill is required to do the sorting, although there is always a need for discrimination when sorting essay type answers.

A major disadvantage is that you need quite a large working space and relative assurance that the piles will not be disturbed. Another is that the task becomes quite involved if the record form is bulky. Turning to the proper page for the sort item, and then to the page for the question you are analyzing, can become irritating and is very time consuming. However, if your record form is a one-page affair, and especially so if it is on a card, you should have no difficulty with this phase of handsorting.

In other respects, the advantages and disadvantages of handsorting are quite similar to those associated with the tally sheet approach to tabulation. Both methods are more subject to error than the other procedures to be discussed. On the other hand, both handsorting and tallying are very practical and effective techniques for determining appropriate classifications and for preliminary reporting of selected data from surveys and other formal studies.

Tabulation Sheet—An elaboration of the tally sheet, and the answer card as well, is the tabulation sheet. This consists of a consolidated listing of answers to all usable questions on all records in the study. The method is most time consuming but there are many opportunities for checking on the work, and you can use the final product for building whatever tables you desire. Unlike the tally sheet, you can find errors more easily since each questionnaire is identified on the sheet.

The setup is much like a tally sheet, with suitable column headings and subheadings. Likewise, you can make your entries with a typewriter or by hand. However, the format differs in that what you would put in the margin or "stub" on a tally sheet is usually included under the major column headings across the top. Each line or "row" across the sheet reports information from a single record form. For an illustration of the format of a simple tabulation sheet, see the selection entitled, "Evaluating a Pasture Improvement Program," appendix A, p. 100.

A great deal of thought must be given to our intended use of the data before we set up a tabulation sheet. Although we do not have to enter all the information we have collected, we should be careful to plan for inclusion of all we are likely to need for our analysis. Once we have planned the format and organization of the tabulation sheet, and completed our tabulation, we cannot refine our data to include more detail. However, if we allow for all eventualities, we can consolidate categories whenever that seems advisable. For this reason, we should make our original entries as detailed as we are likely to have need for them to be. What may seem like unnecessary extravagance of paper and effort at the time may be an economy before the analyses are completed.

In planning our tabulation sheet, we must also consider all of the possible responses to each question and plan our columns and headings accordingly. For instance, where a respondent could answer "yes," "no," or "don't know" to a question, or refuse to answer it entirely, we might properly allow a column for each of these four possible types of responses. This would give us a complete accounting of all the respondents.

In our example of a tabulation sheet (appendix A, p. 100), notice that question #5 has three possible answers: (a) "within past 3 years," (b) "more than 3 years ago," and (c) "don't know" ("D.K."). In addition, we allowed one additional column for nonresponse to the question.

In some instances, one or more questions in a schedule may not apply to some of the respondents. In such cases, a column captioned "NA" ("Not Apply") might be added. Referring again to our example in the appendix, we could quite properly add an "NA" column under question #2. This would then necessitate an entry in the "NA" column for all respondents who had not limed their pasture (e.g., questionnaire #2 and #4 in our illustration).

To separate the data into groups for comparison, the records must be classified **before** tabulation starts. This indicates the necessity of knowing, before you start tabulating, which comparisons you wish to make—what kinds of results you will want. The records can be broken down into any number of subgroups, and the various subgroup totals added together in various combinations to get various group totals. This is, admittedly, a cumbersome method of getting group comparisons and relationships, but it is effective and possible when electric machine tabulation is impossible.

Machine Tabulation—Unless you are interested in quick answers to a few simple questions, and plan to make no intensive study of the answers, machine tabulation generally is the preferred method of tabulation.

However, this method requires access to some very expensive machinery that will not be available to the extension worker for routine evaluations; so we will mention only some of the characteristics of the method in this manual. In any event, if any of you become involved in an evaluation where this method should be used, and facilities can be made available to your research team, you will want to consult with the operators and supervisors of the installation where the data will be processed.

Most land-grant institutions now have such facilities. Extension researchers working on approved projects generally can make arrangements to have their data machine-processed. Persons responsible for the use of the machines should be contacted early in the study.

These people can advise you on setting up your research instruments to secure data most readily processed by machine operations, and they can alert you to the kinds of analysis you can make. They can assist you with coding problems and can show you ways of dealing with free-answer information. They can suggest ways of using machine operations to speed up statistical analyses by providing totals ready for processing in calculating machines. For that matter, some of the more elaborate machines can perform basic statistical operations for you.

The immediate costs of machine processing are sometimes quite impressive, however well-justified. For this reason, there should be a clear understanding of costs before either the researcher or the machine supervisor commits himself beyond recall. They should agree on what operations will be required, and how many. They should agree on the coding, manner of transcribing to machine record forms, amount of checking, number and sequence of sorts and tabulation runs, kind of calculations and correlations to be done, and so on.

There are various companies making different types of machines for tabulating and summarizing answers from record forms or questionnaires. Some questionnaires, set up for certain machines, are marked with a special kind of pencil so that the forms themselves can be run through the machine. The pencil marks make electric contact for a counting of the answers. For other machines, every answer on the record form is reduced to a numerical or alphabetical code; then these codes are punched onto small cards. Coding must be planned by someone who thoroughly understands the study. The cards are run through sorting machines which sorts out the answers to any questions we want, giving the number reporting each type of answer. Or the tabulating machine totals the entries, giving us the sum for all the questionnaires, and the number reporting the answers.

If the questionnaires are set up so that the answers checked are assigned predetermined numerical values (i.e., precoded), the work and time involved in electric machine tabulation is less than for typewriter tabulation, and there is less chance for human error. The great advantage, however, of electric machine tabulation is the increased freedom in manipulation of the cards. Groups do not have to be all predetermined. If an interesting casual relationship is indicated in early summaries of the data, the cards can be sorted on the machines immediately, with no necessity for a complete retabulation as is necessary with typewriter tabulation.

Manual Tabulation Using Punched Cards—Systems have been developed whereby data are coded and punched on cards that are sorted manually. Such a system is useful when the number of cases involved is limited. Extension workers can usually secure help and advice on the use of such cards from college statistical units.

Free-Answer Tabulation—Classification and summarization of answers to open-end questions are difficult. Many times it is impossible to make a usable classification system until the answers are at hand so they can be studied as a unit.

If you have a relatively small sample, you can have all the answers to one question typed out on a single sheet and determine appropriate categories by inspection. What generally proves easier, however, is to enter each answer to the question on a card just large enough to hold the information. Then you start grouping related answers and consolidating related groups until you have a usable number of categories. Now you can assign a code to the categories and use it to make entries on answer cards or other tabulating devices so that you can use these free-answer items as readily as any others in sorting and tabulating.

If you have a large sample, you can save time but still get the job done by going through the same process with a subsample of your records. The first fifty questionnaires might give you the idea. On the other hand, you might prefer to take every "nth" record.

Quite often, in the course of developing categories, you will end up with an "other" or "miscellaneous" category. This will accommodate those responses that do not quite fit into your main classifications or groupings. This will not detract from your analyses unless too many of your records wind up in this "other" group. Any time it becomes your largest category, you should consider redesigning your classification system.

Because so much of the ultimate effectiveness of your study depends on how well you handle your open-end, free-answer, or essay-type questions, it is essential that you make decisions regarding them with the greatest care. This is an occasion to consult with subject-matter and research specialists to be sure that you are making proper decisions. You may even want to set up a panel of experts to work out a classification system for certain items.

This difficulty is the major reason why you should generally restrict the use of free-answer items. Adequate pretesting should suggest ways to set up check-off questions to take their place. Of course, there will be times when elaborate pretesting is not possible or when you are exploring a new topic where you have no idea of what the answers will be. In such cases, free-answer questions are a must.

Another problem lies in the fact that check-off questions limit the range of possible answers. This may contribute to answers which are the result of snap judgment or "any answer to get on with it" reactions. Even if you provide an "other" category, respondents cannot always be counted on to volunteer information. On the other hand, you cannot be sure how complete your replies to free-answer items really are, largely because of this reluctance to volunteer information.

In one case you may have codeable and statistically manipulatable data of uncertain validity; in the other, you have had-to-code, hard-to-analyze, but possibly more valid data. You will have to decide, in terms of the particular item, which type of question will provide the kind of answer that will be most effective for your purposes.

Summary

Our concern in this chapter has been with ways of converting masses of information into understandable and usable form to aid in analysis, reporting, and decision making. We have suggested that tabulation must occupy an important spot in research and evaluation, and that it is most effective when it is planned at the same time the other research instruments are planned.

There are at least four points to consider in planning tabulations: (1) available resources, (2) information needed, (3) relationships of interest to the planners of the study, and (4) comparisons desired. We have suggested that, when in doubt, it is best to collect data in a more detailed form than you are likely to use.

We have described some of the common tabulation techniques, including tallying, handsorting, tabulation sheet, machine processing, and manual tabulation using punched cards. The essential characteristics and the relative advantages and disadvantages of each method are briefly discussed. The chapter concludes with some suggestions on the handling of free-answer responses.

References

1. Croxton, F. E., and Cowden, D. J. **Applied General Statistics.** New York, Prentice-Hall, Inc., 1955, pp. 37-49.

Chapter XVI

ANALYSIS AND INTERPRETATION

Mrs. Jewell G. Fessenden
Ward F. Porter

"All of us must work with figures and numbers every day. The simplest assertion in conversation . . . deals with quantities . . . (Statistical analysis) is one way of bringing big numbers down to a size where they can be grasped with ease."[8]

What we do with figures once they are collected is the key to the usefulness of research findings. Without analysis and interpretation, figures have little or no meaning. Some degree of statistical analysis is within the reach of every extension worker.

Perhaps a definition of the terms analysis and interpretation will be helpful. We can think of **analysis** as the ordering of data in such a manner that they yield answers to questions. This means we take something from our mass of data and organize it so we can take hold of it, catch on to what it is trying to tell us, understand it. We break the whole of our information down into components which become meaningful to us. In layman terms, analysis is putting data through a mill; we take our facts, run them through some sort of statistical processing, and come up with answers. We put in information and take out intelligence.

Interpretation, on the other hand, is the process of searching out the meanings of our answers, of our new intelligence. This process occurs when we get around to answering the "So what?" questions. Interpretation leads to the establishment of explanatory concepts.

In the normal course of events we would expect our interpretations to follow our analysis. It is essential that we have digested (analyzed) data, or knowledge, before we can assign meanings and implications (i.e., make interpretations). However, it is quite possible that our original interpretations may suggest further needs for analysis of our data. This may be the occasion for a new cycle of sorting, analyzing, and interpreting.

When are plans made for the analysis and interpretation of data? This is an area in which much help is needed. A plan for analysis needs to be well thought out by the investigator when a research project is developed. Instruments for collecting data must provide the information needed for purposes of analysis. No amount of analysis can extract from data factors that are not present.

This is not to say, however, that we can make final and complete plans for analysis at the beginning of a research project. The researcher must always look for new concepts as observations are made. However, analysis must precede interpretation, even though there is some overlapping.

It is helpful to write out specific plans and to prepare work tables in advance. This frequently requires a great deal of time on the part of the researcher. It is important that administrators understand that time and other resources must be provided for research. In this connection, a calendar of work, outlining the jobs to be done, the time required, and responsibilities, is helpful.

The processes of analysis and interpretation, even though of fundamental importance in the research process, are often inadequately used by extension workers. This is frequently due to a limited understanding of the basic principles involved. Consideration will now be given to some guiding principles that may be helpful to the inexperienced researcher.

Techniques of Analysis

Statistical analysis can be quite involved. However, all extension workers are capable of using simple mathematical techniques that will serve the purpose for most extension studies. These simple techniques include (1) numerical counts, (2) percentages, (3) measures of central tendency or averages, (4) measures of dispersion, and (5) ranking.

Numerical Counts—For many of our purposes, it may be sufficient to know **how many.** How many of the respondents knew the agent by name (51 did, 17 did not, 2 did not answer)? How many of the respondents were in each age group (5 were under 21 years, 37 were 21 through 54 years, 12 were 55 years or over)? How many of those who were shown how to take a soil sample actually took one on their own land (37 exposed, 13 took sample, 24 did not take sample)?

We need such counts to serve as a base for whatever analysis we do. Sometimes this is all we care to know; at other times this is merely the starting point for an involved series of manipulations. The quality of our data, as well as other factors, will determine how much more we are justified in doing with them by way of analysis.

Caution should be exercised in restricting calculations to simple numerical counts in some cases. Other calculations are frequently helpful in avoiding oversimplification or in promoting clarification. This will become more evident in the discussions to follow.

Percentages—Probably the next most frequently used statistic after the count is the percent, or the proportion the part is of the whole. It is a good aid in making relationships and comparisons meaningful. For example: A relationship like 67 out of 179, compared with 30 out of 70, is more difficult to grasp than 37 percent compared with 43 percent.

Another valid reason for using percentages or ratios is that they are more easily understood. This is particularly true where large numbers are involved. For example, 250 out of 780 does not mean too much to most people; but 32 percent is meaningful.

Although the computation of percentages appears to be a simple and rather obvious process, there are several opportunities for the researcher to go wrong.

1. **Use of the Correct Base**—The base (the denominator or divisor in mathematical terms) is the number from which the percentage is calculated. Selecting the correct base is important. It is equally important to indicate for your readers or listeners what your base is. In a typical evaluation question, the base could be: (1) the total number of respondents in the study; (2) the total number who answered the question; or (3) the total number to whom the question applied.

In using the total number of cases as a base, erroneous conclusions may sometimes be drawn. This is particularly true if the proportion of "no responses" is high. Let us say that out of 100 records, 10 did not respond to a certain question, 70 answered "yes" and 20 answered "no." Using the total of 100 as a base, this means that 70 percent answered "yes."

Using the same figures as above, and taking the number who answered the question at all (90) as a base, we find that 77 percent of those who answered the question reported "yes."

The difference between these two calculations is 7 percent. Which is the better answer? We do not know whether the "no responses" would have been "yes" or "no." Consequently, it is essential that the analysis show either that 10 percent did not answer, or that the 10 "no answers" were omitted.

There are many situations in which a question is not applicable to a respondent. Only the number of persons to whom the particular question applies can be used as the base for calculating a percentage. One example of this is in Table I.

Table I. Broiler practices adopted by 4-H Club members in X county (Number members raising broilers—75)

Practices adopted	Number	Percent
Members attending meeting (where information was given)	50	100
Allowing 1 square foot floor space per bird	30	60
Using thermostatic controls	20	40
Keeping 95 temperature for first week	40	80
Pretesting brooders 24-48 hours prior to use	45	90

This table involves responses to questions asked to determine whether or not 4-H Club members in a broiler project were using recommended practices. In this instance, the county agent is using, as the base, only those who received information on practices. This principle is often applied when an attempt is being made to determine teaching effectiveness.

Another example is the situation where a schedule to be used with a cross-section of farmers has a question on spraying orchards. Obviously, only those farmers in the sample who have orchards will be considered as the base for calculating percentages.

2. **Adding Percentages**—Before we can add percentages, we must be sure that categories are mutually exclusive (do not overlap). For example, let us say we secured the following percentages when respondents indicated how they received information about a particular practice:

Source	Percent
Radio	18
Television	12
Newspaper	28
Neighbors	42
Extension Agent	38
Not known	15

Even if the percentages had failed to add up to more than 100, we should have no difficulty in realizing these are not exclusive categories. It is possible that some respondents could have gained information from more than one source, conceivably from all of the sources mentioned.

However, many questions are such that exclusive categories are possible. This is common among "face" data or background information items. For example, the distribution of replies to a question on age might be shown as follows:

Age group	Number	Percent
Under 30 years	53	9
30 through 50 years	287	47
Over 50 years	266	44
Total	606	100

In this case we have selected our age groups so that there is no overlapping and we are justified in adding up our percentages.

3. **Averaging Percentages**—Sometimes the inexperienced researcher falls into the error of reporting averages of percentages. Although this is a frequently observed practice, it is never quite justified. The following figures are the percentages of 4-H members who completed projects in the different districts of a State:

District	Number	Percent
A	31,000	74
B	8,000	60
C	12,000	65
D	26,000	75
E	11,000	50
F	28,000	72
Total	116,000	66 (incorrect)

In the above table, the analyst incorrectly chose to report, as the total percent, an average of the six district percentages. He should have computed the total percentage figure by dividing the total number of completions (116,000) by the total State membership (165,000) which would result in a figure of approximately 70 percent. Sometimes, as in the above example, the differences are not great. In other instances, the error can be much larger.

4. **Rounding Percentages**—The researcher must decide how many decimal places to include when he reports percentages. To show too many digits may give a false impression of accuracy; to show too few may conceal the fact that differences exist. A safe rule of thumb to employ is to use no more than are necessary to bring out the differences.

Even when your data are precise enough tojustify refined computation of percentages, extra or unneeded decimals add to the difficulty of reading the report and they tend to make the figures harder to remember. As a general compromise, some writers use nearest whole percentage figures in the text of their reports and, wherever necessary, give percentages drawn out to one or more decimal places in the supporting tables or in a footnote.

The following three rules cover the rounding of individual percentages:

1. Where the digit(s) to be dropped represent(s) **more** than half (i.e., more than .5, .50, ½) the numerical distances between the number above and the number below, round to the higher number. For example, the percentages 18.6, 18.51, 17.9 — when rounded to whole numbers — become 19, 19, and 18, respectively.
2. Where the digit(s) to be dropped represent(s) **less** than half of the distance between two numbers, round to the lower number. For example, the percentages 18.4, 18.49, 17.3—when rounded to whole numbers—become 18, 18, and 17, respectively.
3. Where the number to be rounded is exactly halfway between two numbers, round to the nearest **even** number. For example, the percentages 18.5, 18.50, 17.5—when rounded to whole numbers—all become 18.

Most of the time, rounding occurs to the right of the decimal point. Should it occur to the left of the decimal point, zero's must be substituted for each dropped digit. Examples: 169 becomes 170 and 16,987 becomes 17,000.

Measures of Central Tendency — Those devices which help you visualize or identify the central characteristic, the representative unit, or the average, are referred to collectively as measures of central tendency. The ones most likely to be used by the extension researcher are, (1) the mean, (2) the mode, and (3) the median. There are others, but they require some understanding of advanced mathematics and are not meaningful to the larger portion of our probable readers or listeners. Consequently, they have limited utility in extension research analysis and reporting.

The arithmetic mean, or what most people think of when they speak of average, is the most commonly used measure. To get average income, you divide the total income reported by the number of people reporting income. To get average age, you divide the total number of years reported by all respondents by the number of respondents. To get average membership, you divide the total membership by the number of clubs.

The advantages of using the arithmetic mean are: (1) it is most likely to be understood, and (2) it is easy to compute. The disadvantage is that it gives undue value to figures at one end or the other of the distribution. If you were to average club memberships of 5, 9, 9, 11, 13, and 37, your arithmetic mean would be 14, which is larger than all but one of the individual club memberships.

The **mode** is the answer or value that most frequently occurs in a distribution. For example, if more farmers report size of farm as 120 acres than any other size, then 120 is the model size of farm in the study area.

Among its advantages are:

1. It is probably what most people mean when they say "on the average" or "typical."
2. It is the most descriptive average in that it describes the largest number of cases in any one distribution.
3. It is not affected by extreme values. For example, if seven households have the following number of members, respectively—1, 3, 4, 4, 4, 8, 10—the model size family is 4. Regardless of the number of persons in the smallest and largest household, the mode would remain the same.

Its major disadvantage is that it is significant only when a large number of values is available.

Another disadvantage is that occasions may arise when no one group will have numerical superiority over another. Quite often you come up with two groups the same size, or nearly so, and have what is called a **bimodal** distribution. When this occurs, your analysis will be more difficult, but a statistician can tell you how to deal with such situations.

The **median** is the middle value, where half of the cases fall below and half above. To get this value, you arrange (array) your data from one extreme to the other. Then proceed to count half way through the list of numbers to identify your median value. In our first example, the array would read: 5, 9, 9, 11, 13, 37. The number half way through this array would be our median number. In our case, it would be an imaginary number as we have an even number of answers. We can resolve this difficulty by saying our median number lies between 9 and 11, or we can split the difference between these numbers and call it 10. There is a formula for determining this number precisely, but seldom will we need to be this precise.

Advantages of the median are: (1) it is fairly easy to calculate, even when you have to compute the exact median number; and (2) it is not affected by extereme values or the range of data. Its disadvantages lie in the fact that, (1) it is not as familiar as the mean to the layman, and (2) making up the array can take quite a lot of time.

Measures of Dispersion—The frequency distribution is a classification of the answers or values into various groups or classes, arranged in order of size or magnitude.

Classifying people according to place of residence as urban, rural farm, and rural nonfarm is an example of a frequency distribution. Such a distribution might be as follows:

Place of residence	**Number**	**Percent**
Urban	50	10
Rural nonfarm	75	15
Rural farm	375	75

In many cases, measures of central tendency or averages are inadequate for telling the complete story. This is particularly true where the spread or range in the data is great and where extreme values distort the picture. In these instances, the frequency distribution of replies—whether on a numerical or a percentage basis—can be a very useful technique of analysis. For example, it may be desirable to know the proportion of young farm families in a county in order to plan programs to meet their needs.

The use of frequency distributions can be helpful in checking the representativeness of a sample against census data. If this is to be done, the class intervals or classifications should be identical with those used for comparison.

In preparing a frequency distribution, care must be taken to select classes or groups (class intervals) that are meaningful in terms of the local situation. Another requirement is that class intervals should not overlap. In this connection, it would be correct to use "under 10," "10-19," "20-29," etc., rather than "0-10," "10-20," "20-30," etc.

There are, of course, more complex and precise measurements of dispersion than the techniques we have discussed. For the majority of our extension studies, however, numerical and percentage distribution will probably suffice. Whenever we do a study in which our data have been collected in such a manner that more precise analyses are justified, we can consult with a trained statistician who will tell us how to use the more advanced methods.

Ranking—Most people like to know where they stand in relation to their fellows. Clubs like to know how they compare with other clubs on number of members, attendance and participation rates, and other performance factors. County residents like to know where they rank among counties in the State on population, income, percent of farm residents, acres in the different crops, birth and other vital statistics rates, and so on.

Quite often we can use rankings very effectively in our evaluation tasks. Even if the differences may be slight, knowledge that you rank well up on a list can be a source of satisfaction. However, if too much emphasis is placed on the ranking rather than on potential achievements, it can have an inhibiting effect on motivation and incentive to improve.

Rankings are meaningful to most folks, although there is a persistent tendency to interpret rankings as measurements instead of evidences of sequence. It is easy to forget there may be only a small fraction of a unit difference between the 30th and 31st ranked items and a difference of several units between the first and second ranked items, or between the last and next to last items.

Rankings can be made up quite easily if there are only a few items to rank. You just make an array of the items from greatest to smallest in terms of the units of measurements (inches, pounds, acres, members, completions, adoptions). Then you assign consecutive numbers to the items according to their positions in the array. You can assign the numbers from most to least or from least to most, depending on intended use. This point makes no real difference so long as you indicate which way you have chosen to assign the rankings.

Sometimes there is difficulty if several items have the same size (several 5-member families, many 40-acre farms, several members with the same number of project completions, and so on). Frequently, in such cases, it is sufficient to assign the same ranking to each similar entry.

If you are ranking averages, percentages, or index numbers of some kind, you may be able to get differences by carrying out your divisions an extra place or more. This is a situation where differences will be trifling and the resulting rankings may have little meaning.

If there is no way to break the tie, or if you feel there is no justification for assigning different rankings, you give all equal items the same ranking. Then you must be on the alert to avoid a very common error in ranking. You must leave places in your ranking for each tied entry and assign the next entry the first place beyond these.

If your array reads, G-9, A-8, D-8, H-8, B-6, E-5, C-3, F-3, you would report rankings, G first; A, D and H tied for second; B fifth; E sixth; and C and F tied for seventh; not, G first; A, D and H tied for second; B third; E fourth; and C and F tied for fifth.

Rankings are not used a great deal in statistical analysis beyond simply reporting them, although it is possible to use them to make rough correlation tests. Because of some problems in dealing with ties, however, it is best to ask a statistician to explain the formula and the limitations of the statistic to you.

Interpretation Techniques

After we have collected, tabulated and analyzed our data through the various processes available to us, we still have the task of interpreting the findings. It is not enough merely to cite our findings in descriptive fashion. We must ask ourselves the question, "What do these data really mean?" More than that, we have to come up with the answer to the question, or at least a statement of what we believe the findings imply. To guide us in working out such a statement, there are several fundamental principles which we will now consider.

Generalization—Most research is motivated by a desire to make inclusive statements based on a limited number of observations. We desire to observe results in a limited setting and generalize our findings to a larger group or population. The demands such a desire places on our methodology has been discussed at considerable length in other chapters.

We must exercise caution in drawing conclusions for a larger or different population than the one from which the sample was drawn. If we collect information from only farm residents in the community, we cannot generalize to the total population of the community. In fact, if the percentage of returns is low (as in many mail questionnaire studies), or if the refusal rate is high, we may be justified in applying the findings only to the actual group of respondents, rather than to the whole of the sample and to the population from which the sample was drawn, as we had intended.

Definition—We have no right to generalize for areas or situations where the definitions we have used in our study do not apply. If we would make comparisons, in terms of concepts or standards or characteristics that are not uniformly defined, we must attempt to either adjust our findings so they will be comparable, or acknowledge that comparisons cannot be made, or that they are made with reservations and qualifications.

We may want to compare our findings regarding income in home demonstration clubs with census data. Unless our definition of income is the same as that of the census, or unless we can go back to the original data and work out comparable units, we are not justified in making income comparisons. If the comparison is to be valid, incomes must be defined in the same way.

Objectivity—Again we must remind ourselves to keep personal judgments and other evidences of subjectivity out of our interpretations. Now is the time when the attention we gave to assuring objectivity in setting up our record forms will pay off.

For example, "Did you eat an adequate breakfast this morning?" leaves the interpretation of "adequate" to the person who answers the question. "What did you eat for breakfast?" requires the listing of all foods eaten. The adequacy of the breakfast can then be determined in line with recommended nutritional standards.

Cause and Effect Relationships—One of the basic purposes of scientific research is to establish causality (the reason why). Now we are all aware of how difficult this is to accomplish in the behavioral sciences. However, our scientific inquisitiveness compels us to try to do it, and we feel certain that as our techniques improve we will become increasingly successful in doing it. The best safeguard we can employ before making claims for cause and effect relationships is to screen out all other likely causal factors before drawing our conclusions.

In the table that follows we have what looks like a good case for the value of extension teaching. This is an illustration from an extension study designed to learn the effect of participation in Extension on the rate of adoption of recommended orchard practices. Strictly on a basis of percentage comparisons between participators and nonparticipators, we appear to have confirmation of the effectiveness of participation in Extension as a causal factor in the adoption of recommended orchard practices.

Table II. Farm practices carried out with relation to farm operators' participation in Extension.

Practices adopted*	Extent of Participation		
	Active	Inactive	None
Percentage of those with orchards who:	— — percent — —		
Carried out 3 or more recommended practices	76	68	50
Applied fertilizer	63	78	36
Cultivated adequately	75	80	82
Followed spray schedule	30	20	—
Treated for borers	32	20	18
Pruned trees	68	53	36
Thinned fruit	26	23	20

* All practices were defined.

However, there were some other differences among the farmers in the sample, in addition to extent of participation. We found that the more active participants tended to have higher education levels, higher income levels, and larger farm operations. Even though more refined statistical calculations might reveal a positive relationship, there is room to challenge our unqualified acceptance of extension teaching as the sole, or at least major, causal factor. It now appears obvious that some of these other factors may be contributing factors in the adoption of recommended orchard practices.

Our point here is that we must be extremely careful whenever we assign cause and effect relationships. Before we make such claims we must be sure to check our relationships against all available evidence. This means drawing on our "face" data items to make all likely cross-tabulations to identify any associations that may supplement our basic relationship in any way. This is also a final plea to include adequate "face" data and control items in all data-collecting forms.

Program Objectives—Many of our studies are made to determine program effectiveness. Data collected for this purpose might be interpreted to point out how well program objectives are being met. These objectives, as suggested elsewhere, are expressed most aptly in terms of specific behavior changes.

In the table that follows, the desired behavior change is **constant** feeding of phenothiazine and salt. You will note that close to three out of every ten of the farms made the desired behavior change, almost a half partially adopted the practice, and about a fourth failed to act in the desired manner. The effectiveness of the teaching will be judged in accordance with what is considered a satisfactory rate of change for the particular objective.

Table III. Use of phenothiazine and salt on farms in sample, X county, 1950.

Degree of use	Number of farms*	Percent of farms*
Feeding constantly	35	29
Feeding occasionally	55	47
Not feeding at all	28	24
Total	118	100

* Selling sheep and lambs.

Potential—Much of our interpretation is concerned with determining how successful we have been in reaching our potential clientele. Potentials are closely associated with program objectives and with careful definitions of our population in terms of them. In 4-H Club work, for instance, the objective may be to enroll boys and girls who are students in rural schools. In this case, the potential would be all the boys and girls of 4-H Club age who are enrolled in rural schools of the area under consideration.

When it comes to interpreting enrollment figures on a percentage basis, only those boys and girls of the proper age and going to rural schools would be considered.

On the other hand, if all boys and girls of 4-H Club age who live in the county are encouraged to enroll in 4-H Club work, then the total number of such young people in the county would be the potential clientele for the action program. This number would serve as the base for computing the percentage of the potential reached.

Another example of interpretation in terms of potentials relates to pasture improvement. One objective of such a program might be the systematic application of lime. In this instance, our potential audience is limited to those farms having pasture land. Secondly, we would consider only those pastures with soil conditions requiring lime. In other words, we would not have a true picture of the effectiveness of our teaching if we included all pasture acres, and less so if we included all acres in farms.

Trends—We do have a good deal of trend analysis and interpretation in extension evaluation work. Before, during and after data are often the basic of our evaluations of extension teaching effectiveness. Any time we do a benchmark study we are setting the stage for a trend study. Trends are very helpful in any study of change, whether as indicators of progress or direction of movement.

We have to be very careful when interpreting changes, however, as there are some factors which could be involved besides the ones under consideration. Frequently, what appears to be a marked change may be accounted for almost entirely by a shift in population.

For instance, we have the example of the report by presumably qualified researchers stating that there had been a 10 percent increase in milk consumption in their State during a specified period, something like a shift from 2,000,000 pounds to 2,200,000 pounds per day in terms of whole milk equivalent. Actually, a realistic interpretation of their findings would have indicated virtually no change in milk consumption per capita during the period. The population increased from something like 2,000,000 to about 2,200,000 persons during the same period; so we can see that instead of increasing, per capita consumption remained almost constant at approximately one pound per day.

We are on even more dangerous ground when we try to interpret trend data in terms of changes in dollar income over a period of years. Inflationary tendencies in the economy as a whole may account for any apparent increase in individual incomes.

When you find yourself involved in such interpretations, consultation with a trained researcher or a specialist in statistics is in order.

Summary

After some preliminary comments to identify analysis as the ordering of data to get answers to questions, and interpretation as the process of searching out meanings of answers, this chapter points out some principles and techniques of analysis and interpretation.

In our discussion of analysis techniques we refer to counts, percentages, measures of central tendency and dispersion, and ranking. Those techniques requiring only an elementary knowledge of mathematics are emphasized. We also point out some common errors in the use of each technique.

In discussing interpretation techniques, some safeguards are suggested to assure objectivity and appropriateness in our appraisals. We mention some factors limiting how far we can generalize. We stress the importance of agreement on definitions when making comparisons. We point out the need for adequate checking before postulating cause and effect relationships. Finally, we discuss the importance and value of interpretations in the light of potentials, and trends.

References

1. Croxton, F. E., and Cowden, D. J. **Applied General Statistics.** New York, Prentice-Hall, Inc., 1955, Chapters 1, 7, 8 and 9.
2. Good, C. V.; Scates, D. E. **Methods of Research.** New York, Appleton-Century-Crofts, Inc., 1954.
3. Royal Bank of Canada. "The Uses of Statistics," **Monthly Letter,** January 1956.

Chapter XVII

PREPARING A RESEARCH OR STUDY REPORT

Mrs. Jewell G. Fessenden
Ward F. Porter

If research findings are to be used, they must be organized and made available in understandable and appropriate forms. In some cases a complete written report is not required. If the results of a study are to be used only by the investigator, or a few people, there may be no need for a report other than analysis tables. Examples of this may include a summary of personnel ratings or a determination of interests of a specific group with reference to conferences, seminars and related activities.

In this case there may be no distribution of written materials. In general, however, if research results are to be used by many people, a written report will be necessary. Sometimes when the study is an important one, a single report will not do justice to the findings. In this case, a series of reports, aimed at different audiences, may need to be prepared. The principles of presenting findings would apply almost equally to both the oral and written presentations.

Data are collected to answer questions. When interpreted, they supply answers to those questions. The report states the questions, outlines the method of collection and analysis of the data, and explains the answers obtained within the limitations of the study objectives and findings.

In preparing study reports, there are many factors to be considered. This chapter is concerned with some of the more important considerations. Among these, we will discuss the anticipated audience, the report outline, contents and organization, writing styles, and methods of presenting study data. In addition to these, one would also need to consider available resources, pertinent to preparing and publishing reports. These aspects of report writing will be discussed under the three major headings of planning, style of writing, and ways of presenting data.

PLANNING A STUDY REPORT

The Audience

Any writer who hopes to do an effective report should have some idea of his probable audience. Some understanding of its needs, interests and background will help him decide which points to stress in his presentation. With this understanding, he will be in a better position to know when to slow down and explain things; he will know which details to mention briefly and which ones to spell out; he will know the proper degree of familiarity to use in his writing style.

For our present purposes, we may conceive of three types of audiences (readers or users) for extension study reports: (1) other research people; (2) all professional extension workers and other educational personnel in related fields; (3) the lay audience. In any one of these categories there may be several subclasses. For example, the lay audience as thought of in this chapter may consist of rural, urban, farm, nonfarm or other people not especially concerned with professional extension work.

If a report is prepared for **other research people,** it is important to include the following:

1. Table of contents.
2. Summary statement of findings.
3. Specific purposes and study objectives.
4. Significance of the study. This should include reasons for undertaking the project and its meaning to the audience concerned.
5. Procedures used in the study. This might involve a complete discussion of: population and sampling; methods of data collection—mail, personal interview, telephone, or combinations of these methods, record forms used; selection and training of interviewers; the dates on which field data were collected; and rate of responses.
6. Statistical methods and techniques used in the tabulation, analysis and interpretation of the data.
7. Results obtained in relation to the objectives of the study as well as other findings of a "byproduct" nature.
8. Implications of the findings in relation to the purposes of the specific study or research project. This may be less important for research

persons than for the other audiences previously mentioned.

9. Identification, including the names of people responsible for planning and conducting the study, date and place of publication, acknowledgments.
10. Reproduction of survey schedules or questionnaires.

The order in which these phases of the report are presented is sometimes determined by the author. In other instances, the institution or agency sponsoring the study makes the decision.

If the audience is expected to be **professional extension and other educational leaders,** certain adaptations of the above content, organization and format of the report may need to be made. For this group, details as to statistical procedures, sampling techniques and other methods used in conducting the study may be omitted or at least discussed only briefly. It is important to report the findings in such a way that the extension worker can see their application to his own program. This may be the improvement of a method, planning procedures, or program content.

For the lay audience, a research report needs to be concise, non-technical, easy to read, and may contain only major findings from the study. Raising questions relating to the significance of the research for the particular audience may make a report more personal, appealing, and, therefore, meaningful.

The above discussion implies that there may be more than one report prepared to meet the needs of various audiences. In some cases there may be several reports. For example, there may be a very complete, detailed and highly technical report for the audience of **research personnel;** a report for **professional extension workers** containing the major findings and implications, with less emphasis on research procedures and technical details; and, finally, a relatively simple and brief report for **lay people** highlighting the major findings in a popular style.

It is also appropriate to think of a highly comprehensive study being reported in several specialized "editions." For example, a survey may show certain data on topics of interest to different specialists.

Resources

Some other factors to be considered in planning a research report include resources available to the researcher and the institution or agency involved. Among these are time, finances, facilities and personnel.

Time is of importance from at least two standpoints; that of time available to prepare the report, and the urgency in terms of timeliness and use of the findings. Adequate time should be made available for the preparation of the research report. It is wasteful of time and money to make an investigation without adequately reporting the results.

The money, facilities and skills available for writing and illustrating will influence the type, number and scope of research reports. This should also be taken into account when the research project or evaluation is planned.

It is suggested that the author of a report work very closely with the editorial staff in charge of publications, both in planning and writing a report. Editors can be of invaluable help in selection of visuals, improving readability, and effective presentation.

WRITING A STUDY REPORT

The Report Outline

"In planning the report, try to think through the story that you want to tell your readers; then write that story in your own words. Report writing is not too different from other writing."[2] Effective writing is a tool that helps to insure understanding and use of study findings.

Most writers who contemplate writing a report find it helpful to have some sort of outline to work with. It will assure order in the finished work and it will help hold down repetition and guard against omissions. Such an outline need not be spelled out in complete detail and does not always have to be written, but the various points to be included should be given careful thought before actual writing is started.

Organization and Content

Select only the facts that are important and significant for the various audiences for whom the report is being written. It is helpful to arrange facts in logical order so that they may be easily followed by the reader. The use of subheads helps to call attention to points that need to be emphasized and facilitates readability.

Findings should be reported in terms of the study objectives. Wherever results are not conclusive, some explanation should be made. It is also important that an objective approach be maintained in interpreting the findings. Where the writer is expressing a value judgment, this should be made apparent. For example, a writer may state that the rate of adoption of a practice was "high" when 25 percent of the respondents reported use of the practice. In reality 25 percent may or may not be a "high" rate, depending on the circumstances or previously established standards.

There are several important considerations regarding the organization and content of a study report.

1. **Title**—All reports should start off with a title, either on a separate page or at the head of the article. In the more popular types of report it should be short and designed to attract attention yet give some clue to contents; in the more formal reports it can be longer and should give a more positive clue to contents.

Most of the time the author's name appears immediately after the title or somewhere on the title page. Sometimes his name does not appear on the more popular types of report and his agency is designated as the author.

2. **Summary Statement of Study**—An increasingly popular practice in report writing is to place a summary of the report in the fore part of the publication. It may come before or after the table of contents and list of charts, figures and tables, but it always precedes the body of the report. It is not designed to take the place of the chapter on conclusions which continues to be placed at the end. Rather, it pulls together the highlights of the entire report and presents them so that, (1) the mildly interested reader can get the main idea without reading further, and (2) the undecided reader can get enough information to decide whether or not to read the complete report.

3. **Table of Contents**—Quite often the table of contents is omitted from short reports and those of intermediate length which are written for the lay audience. Any of the longer reports, regardless of the intended audience, need a table of contents to speed up reference or to locate chapters or sections in which the reader may be most interested. In the very detailed report, the table of contents may list subtitles as well as chapter and section titles.

4. **Study Objectives and Procedures**—Any report, whether it is a professional monograph, a graduate student thesis, a journal article or note, a miscellaneous publication, or a leaflet, requires an introductory section of some sort.

One of the items to be covered in such a section is a statement of the objectives or purposes of the study. This statement should also contain the reasons why the objectives were considered important, and the probable use that could be made of the findings of the study.

Somewhere in the introductory section there should be some comments on how the study was made. The amount of space to devote to this will depend in large part on the intended audience. We would not expect certain lay people to be interested in all of the details, although it is a good idea to give them some notion of where the study was made, who was involved, and procedures used in collecting the information.

The educators would want a little more background information to aid them in assessing the merits of the study.

The details on procedures will be of major interest to other researchers. They will want to know our population and the sampling plan. They will want to know what kind of data collecting instruments we used, and they would like to find a copy of each in the appendix section. They will expect to find some notes identifying the research team and what training or conditioning was provided for them. They will be interested in details like the inclusive dates of the field work, any special techniques that were used during the field work phase, the response rate, and so on.

5. **Findings**—When you have completed your introductory section, you start reporting your findings, sometimes with the help of tables, charts, graphs and figures, or reference to appropriate appendix items.

You should report your results or findings in terms of your study objectives and should present the evidence both for and against your hypotheses. When you feel you should express a person opinion, you should say so. Whenever results are not conclusive, some explanation should be made. You should also feel obligated to mention any findings of a byproduct nature which you consider to be contributions to knowledge or understanding anywhere in the general area of your study. The limitations of the study should also be pointed out, in all fairness to the reader.

6. **Implications**—After you have presented your findings, you have an obligation to make some comments as to what the findings mean. At this point it is extremely important to involve subject matter specialists in interpreting the findings and in suggesting implications. Now is the occasion for suggesting related topics for future research and for pointing out facets of your own study which should be examined more closely, possibly with new data.

The need to include implications in your report is most urgent when you are writing for your lay audience. This audience is least likely to recognize the possibilities of your findings, partly because of their less thorough knowledge of the subject, but also because of less familiarity with dealing in cause and effect relationships.

Most members of your educator audience will also need some assistance with implications. In addition to your implications, they will draw their own after they have considered the report in relation to their own experiences. Because they do this, you have the additional responsibility of being certain you make no implications which your data do not justify.

7. **Documentation**—Most reports of research require a certain amount of substantiation from sources other than the research data at hand. Sometimes, in your report, you may want to make comparisons of your results with those obtained from other sources. These other sources should be identified by footnotes or notes at the end of the chapter or section. All of the sources may be assembled in a bibliography to accompany the report, either as part of or prior to the appendix.

The bibliography and footnotes frequently are omitted from reports for the lay audience but they should be in reports intended for educators and researchers.

8. **Appendix**—This section serves as the catchall for items which had to be omitted from the body of the report for one reason or another but which were important enough to be included with the report. Sometimes side issues are elaborated on in separate appendix notes. This is where you put your working tables which are not needed for supporting the test but which might be of considerable interest to some of your readers.

Here you can elaborate on your sampling plan or demonstrate how to derive the statistical formulae you employed. You can include extra maps, charts, record forms used in data collecting, and any devices you may have worked up as a result of your findings.

Although used extensively in scientific research reporting, appendix sections are not common in popular reports.

Style of Writing

Effective writing is a tool that helps assure both understanding and use of study findings. Mrs. Cowing says that the specialist cannot hope to write "best seller" articles while reporting research findings, but he can make his report more readable. She offers this advice:

> "You technical writers cannot assume that lay leaders will be interested in your technical information. You have to identify your technical subject with their interests; let them know it will pay them to read it. Slant your writings to their interests as well as to their needs. Pack your writing with appeal; tell them it's "good-to-eat" as well as "good-for-you." Camouflage what you think they ought to read with what they want to read."[3]

Although there are occasions in extension research reporting when "professional" writing styles are in order, most of the writing we do in Extension will be on the "popular" side. For this reason, we will devote our remarks to ways of improving our writing style to make it more appealing to our educators and lay audiences.

Some characteristics of effective popular writing are to: (1) use the personal approach; (2) use short sentences whenever possible; (3) use the simplest word that will carry your intended meaning, and be sure a long or technical word is understood when have to use it (4) confine your remarks to the main facts, and keep them flowing in a logical order; (5) hold down the number of qualifying statements, statistical references and parenthetical comments; (6) use the simplest illustrations or examples that will accomplish your purpose; and (7) use present tense and active verbs.

Most of us object to rewriting, but that is the best way to improve the readability of our writing efforts. If popular writing does not come naturally to you, just write your report to include all the points you want to make, using whatever style you like. Then go back over it and break up those long paragraphs into shorter ones containing a single idea.

Go through again and change from the less personal to the more personal pronouns. Next, substitute shorter words for as many of the long words as you can. When you must use technical words, define each one the first time it appears. Take a second look at any sentence over three typewritten lines long. See if you can make it shorter, or rewrite it as several short sentences.

You may have to go back and regroup some of your paragraphs and this may call for some different transition phrases or sentences. Work in a few short subheadings and see if any words or phrases in sentences need underlining for emphasis. Type up a clean copy, and see how easy it is to read and understand!

METHODS OF PRESENTING DATA

There are several things in addition to writing style that an author must consider while preparing his report. None of our audiences would be pleased with a report consisting of narrative description only. They like the presentation to contain more than the printed word. We will discuss some of the more favored ways of presenting research data.

Narrative Description

The most common form of presentation, either alone or supplemented by other forms, is the narrative report of findings, accompanied by whatever conclusions, implications and recommendations the author cares to make. The narrative statement alone can become very boring, but it is appreciated by those who do not care to puzzle through tables or charts.

Tables

We use tables to bring together in rows and columns larger amounts of data than we would care to describe in words. We are still at liberty to make whatever comments we want to about the data, but we do not have to include so much detail when we provide tables containing all relevant data. Large tables, like Table I, generally are assigned to the appendix section; smaller ones, like the next two, are more apt to follow along with the narrative development of the report.

Table I. Percentage of Respondents 14 Years Old and Over Who Had Completed a Specific Number of Grades in School, by Age, Sex, and Industry Group, Upper Monongahela Valley, West Virginia, July 1954.

Group	Total persons reporting		Percentage of persons whose highest grade completed was						
			Elementary school			High school		College	
			0-3 Yrs.	4-7 Yrs.	8 Yrs.	1-3 Yrs.	4 Yrs.	1-3 Yrs.	4 or more
	No.	%	%	%	%	%	%	%	%
All persons	2,252*	100	6	22	30	20	16	4	2
By age and sex									
Male	1,114	100	8	23	31	18	14	4	2
14 - 24	261	100	2	21	17	37	19	4	**
25 - 64	68[illegible]	100	7	21	36	14	14	5	3
65 and over	16[illegible]	100	21	35	36	3	3	1	1
Female	1,138	100	5	21	29	21	18	4	2
14 - 24	267	100	1	9	15	46	24	4	1
25 - 64	726	100	4	22	33	16	19	4	2
65 and over	145	100	16	39	32	6	5	1	1
By chief income source									
Agriculture	190	100	9	22	39	12	12	4	2
Mining	458	100	6	23	33	23	13	2	**
Manufacturing	475	100	6	20	28	23	19	3	1
Business and professional	211	100	2	16	23	25	23	9	2
Other nonfarm	398	100	2	12	29	20	24	7	6
Nonwork***	396	100	15	37	28	11	7	1	1
Other and no report	124	100	3	22	31	23	15	3	3

* 35 persons did not report.

** Less than 0.5 percent.

*** Nonwork households are those whose major income was from nonwork sources, e.g., rents, royalties, Social Security, retirement funds.

The second table is one designed to illustrate a few simple comparisons. It may not involve all respondents and may not exhaust the information provided by a single question, but it presents the data to support the narrative at the particular point in the report. It requires a title, headings and a few guidelines, but may not have subheadings or involved breakdowns in the stub.

Table II. Ideas Used in Teaching Economy Meals to Members and Nonmembers

Subject	Percent	
	Members	Nonmembers
Knowing when one food is a better buy than another	80	49
Knowing what to look for in reading labels	79	42
Knowing how to make substitutions in market list without changing food values	77	35
Meal planning	71	35
Making a market list	69	35
Using basic seven foods chart	68	40

The third table is a kind which is used quite often to point up some simple relationship. This kind is included in the text, does not carry a number or title and usually contains only one column of figures.

Reading and studying	20%
Sewing	12%
Family recreational activities	10%
Laundry work	42%

Graphs and Charts

There are many kinds of charts, graphs, maps, sketches and diagrams that can be used to present data. Compared with tabular methods, graphs and charts have some advantages as well as limitations. They are more likely to attract attention and impress the facts upon the minds of the readers. They are also more limited as to the quantity of data that can be presented effectively. A single chart or graph should not be complicated or attempt to present too many facts which might be confusing. Unlike tables, graphic methods can usually show only approximate values.

Line charts consist of broken or curved lines that are used to picture a time series or trend and frequency distributions. Care must be used in selecting the intervals on both the horizontal axis (baseline) and the vertical axis. False visual impressions can be created by using intervals that are either too broad or too narrow for the data in question. Curves or line charts are not appropriate where the items or units represented are discreet or where it is incorrect to read off values that lie between the intervals on the baseline. In such cases, a bar chart is to be preferred.

Bar charts are made up of horizontal or vertical bars and are used mostly for comparing magnitudes. The difference in length of bars on a chart is, for most people, easier to see than the difference between numerical figures. Bars are also used advantageously for showing component parts, that is, the parts of a total.

Pie charts have a wide appeal and are particularly effective in indicating the whole and component parts. Greater accuracy in reading pie charts is insured if numerals are written on the various parts of the graph.

Pictorial charts lend variety and attract attention. However, they should not be used to show too many ideas at once. The main rule to remember is that an increase in magnitude is indicated by an increase in the number of characters and not by an increase in the size of the pictures or symbols! Symbols should be self-explanatory.

Summary

Research has little value if it is never pulled togetner into some form of report. Although an oral report may suffice for some occasions, we generally think of the written report when we think of reporting. To guide the author in preparing the report, we mention several things which can simplify and give direction to his efforts.

The first point to consider is the probable or intended audience. We identify three major classes of audience, the researcher, educator and lay audience, and give some of the core characteristics of each.

The next item we consider is the outline and organization of the report. Its major elements are the title; summary statement; tables of contents; objectives and procedures; findings; implications; documentation; and appendix.

In our discussion of style in writing the report, we stress the techniques for achieving easier and more readable manuscripts. Most of these involve simplicity, brevity and organization, and are most efficiently accomplished by conscientious rewriting with the objectives of effective popular writing clearly before us.

Our final topic is methods of presenting data. We suggest the use of and briefly describe illustrative materials, such as tables, charts, maps, graphs and pictures.

In the final analysis, decisions on writing style and methods of presentation must depend on the intended purposes of the report and on its prospective readers. A report must tell the story effectively, but this is not enough. There must be a personal appeal that makes the reader say to himself, "This report has something in it for me."

References

1. Clark, R. M., and Norford, C. **Materials and Techniques for Making Charts and Graphs in Teaching Agriculture.** East Lansing, Vocational Education Department, School of Education, Michigan State University, 1954.
2. Cowing, A. G. **Let's Write Easy Reading.** U. S. Dept. of Agr., Fed. Ext. Serv. Mimeo, 1111 (11-47), Washington, D. C., 1947.
3. Cowing, A. G. **How to Communicate Ideas in Words—Plan, Write and Trim.** Unpublished talk given at the Clear Writing Workshop, held by the Soil Conservation Service, U. S. Dept. of Agr., Washington, D. C., May, 1956.
4. Flesch, R. "How to Say It with Statistics." Reprinted from **Printers' Ink** by Division of Extension Research and Training, Fed. Ext. Serv., U. S. Dept. of Agr., Mimeo. 248 (3-51).
5. Good, C. V., and Scates, D. E. **Methods of Research.** New York, Appleton-Century-Crofts, Inc., 1954.
6. Parten, M. B. **Surveys, Polls and Samples.** New York, Harper and Bros., 1950.
7. Strang, R. "Principles of Readability Applied to Reporting Research," **Improving Educational Research.** Washington, D. C., American Educational Research Association, 1948.
8. U. S. Dept. of Agriculture, Office of Information. Examples of Charts Used in Presenting Statistical Data, Negative 45145.
9. Yang, E. H. P. **Fact-Finding with Rural People.** New York, International Documents Service, FAO, 1955.

THE USES OF EVALUATION RESULTS

Chapter XVIII

APPLYING THE RESULTS OF EVALUATION IN EXTENSION

Joseph L. Matthews

Four Objectives

The objectives of this chapter are to help you:

1. Appreciate the potential for gain that can come from using results of evaluation in your extension work.
2. Appreciate the importance of evaluation procedures as aids to getting results incorporated into action plans.
3. Learn ways of getting results from extension evaluation and other social science research used.
4. Be motivated to use results from extension evaluation in your own work.

Three Ways to Improve Practices

We can use three ways to improve our extension educational practices. The easiest way, and the one least likely to provide the best solution, is to make changes on the basis of our **subjective impressions** of what the problems are. With a minimum of testing or accumulation of objective evidence, we can make judgments about probable consequences or outcomes of our contemplated improvements, and be guided accordingly.

Another relatively easy way is to ask a **professional research** person to study the problem and suggest solutions. We canot be criticized for using this way on grounds of unreliability of results or defective methodology, but we can be criticized for passing up the learning experience of finding out for ourselves. At any rate, there will be few occasions when we can afford to have a professional do the kind of job that is needed. More than likely we would have to settle for his stock answers based on reasonably similar situations reported in the literature.

Somewhere in the middle ground between these two methods lies the way which is the subject of this manual. **Evaluation** combines the advantages of first-hand association inherent in the first way, with some of the more manageable objective and statistically sound techniques of the latter, into a workable plan for use by any reasonably competent extension worker.

The final product resembles traditional research in most fundamental ways. In each there is an attempt to define the problem with precision, to derive the hypotheses from the best available background information, to follow methods that will provide facts rather than subjective impressions, and to generalize from the evidence with caution.

Basic Differences

The one basic difference centers on the matter of the **motivation** of persons making the study. The typical research person is motivated by a desire to arrive at "truth." In his thinking, the degree of excellence of the work depends on the degree to which it contributes to establishing generalizations that have wide applicability.

On the other hand, those of us who are involved in evaluating our own efforts are motivated by a desire to improve our own ways of doing our work. When such is our objective, coming up with answers that have relatively unrestricted application in Extension are of little immediate concern. We want answers that apply to our own particular working situation.

This kind of purpose does not mean that our approach to evaluation should be one that restricts application of our findings to a particular situation. Quite the contrary! The sound research methods we apply while evaluating a particular situation should produce clues or suggest hypotheses that can, in some degree, be useful for analyzing a similar situation in other settings, as well as our own. We owe it to our fellows to do the kind of evaluation job that will make our findings as meaningful and useful as possible to them, so long as they first answer our own questions.

In addition to fundamental orientation, the **working conditions** under which the two types of investigators carry out their inquiries may differ appreciably. Traditional researchers in education attempt to control the study setting so that many of the variables normally involved in a real situation are ruled out or controlled in some way by the study design. By so doing, they gain precision at the expense of relevance of findings.

Extension workers who engage in evaluation must conduct their inquiries in the complicated climate of their regular activities. Because of the large number of variables in such a setting that cannot be controlled, the resulting evaluation may be lacking in precision. Nevertheless, results so arrived at have real meaning because they are derived from an inquiry carried out in a real situation. Any solutions suggested by such findings would not have to be translated from the laboratory setting to the real one.

Disadvantages of Calling in Professional Researchers

Relying on the professional research person to take care of our evaluation problems has two basic limitations. We have already suggested the first one. Such a person cannot study individual problems in any strict sense of the word because his orientation is toward generalization.

Impersonal—He may be able to think of your problem as an opportunity to try out a pilot study, in which event he can personalize the study enough to provide the kind of answers you need. In most cases, however, he must design studies which promise to contribute to general rather than specific knowledge. When most of our problems are peculiar to a situation and to the people who are involved in it, we cannot hope to gain maximum benefit from the kind of study he is prepared to make.

General—A second limitation is that, even when the professional researcher is able to make sound recommendations, we will have difficulty incorporating them into our day-to-day activities. It is easy to talk about a better way of doing something as a result of reading about or being told about what others say should be done. However, there is a great difference in likelihood of acceptance between change recommended by others, however qualified to make recommendations, and change suggested by our own careful examination of the situation.

Detached—There is a wide gap between talking change and accomplishing change. The surest way to bridge the gap is for us to participate actively in decision making as to what the change should be so that we thoroughly understand the matter and do not have to waste time convincing ourselves before we can hope to convince others. Then we can proceed to convince others by example as well as talk.

We can expect our associates and clientele to react much more readily to the suggestion and example of someone they know personally, such as ourselves, than to the most positively uttered recommendations of the professional researcher. Such a person cannot afford to take the time to identify properly with our group. He has to rely for rapport on his professional title, the prestige of his sponsoring agency, or whatever aura of authority we can build up for him.

Getting Results

This suggests the next point we might consider: "How can we go about getting the results from evaluation used?" A related question is, "Who can use the findings from evaluation?" Obviously, the only ones who can use the results of evaluation are the persons who are involved in the situation that was studied or in another situation that is similar.

The next obvious thought is that, regardless of who is involved and regardless of the situation, the results must be made known to the potential users. The quality of the knowing is critical. Adequate means of involvement must be devised if the knowledge is to be used to fullest advantage.

We use **involvement** here in the sense of bringing appropriate persons into intimate contact with results of the evaluation. Ideally, this is achieved most fully when the persons in a position to use the results actually do the evaluation.

The problem is little different from getting farmers to use the results of scientific research in agricultural production. The difference, in the main, is in the matter of motivation and the consequences of not using. When the advantages are not readily evident, or when the penalties for noncompliance are not immediately apparent, it is much harder to convince anyone that action is necessary.

Involving Others in Evaluation

As more people become involved in an evaluation, the difficulty of **fully** involving all of them increases out of proportion to increase in numbers. Although we cannot hope to keep all of them involved completely all of the time, we can work out ways to keep as many of them as possible involved as completely as possible. We will discuss things that we can do to this end in terms of five stages and the main questions to be answered at each.

We must realize that some of these suggestions will not apply for all evaluation efforts, but most of them will have quite general application. Many evaluation efforts are concerned with the work of a single person and may be related only to a single part of his work. In such evaluations, there may be little or no need to involve others. Nevertheless, there are many occasions when it is best to share with others the results of even the simplest evaluations. For this reason, most of the following will apply in some degree to all levels of evaluation.

1. **When Deciding to Evaluate**—Two questions come to mind at this stage: "Shall we evaluate?" and "Exactly what shall we evaluate?" Evaluation should be undertaken in response to an expressed desire on the part of the potential users of the results.

Sometimes we may have to take it upon ourselves to bring the need for evaluation to the attention of our associates or clientele and to kindle a desire in them to have the evaluation take place. Frequently, just showing how to evaluate is enough to stimulate a desire to try it. We can do this ourselves or bring in someone from outside the group to do it for us. This can lead to interesting them in reading and reporting on findings of evaluations or extension research carried out elsewhere, with the idea of deciding whether or not evaluation is needed locally.

Once the idea of undertaking an evaluation begins receiving serious consideration, the various members can be kept involved by encouraging them to contribute ideas for the evaluation itself and for how they might be able to make use of results in their own work. This can be done with each individual in separate conferences, in meetings of the entire group, and in committee work. Action should include consideration of possible findings and the changes that might be in order as a consequence. Discussion should include possible alternative courses of action and their probable consequences, constructive or otherwise.

2. **While Planning the Evaluation** — The next phase is focused by the question, "How shall we do the evaluation?" Now is the time to crystallize thinking about extension objectives related to the particular topic by research, discussion and formal statements of understanding. This is also an excellent occasion for us to remind our associates and clientele of related basic extension objectives and ideals.

This is when available alternatives must be considered to decide what kind and quality of evaluation to do. Thorough discussion of methodology and what can be expected in the way of results from the various methods will contribute to better understood decisions regarding budgetary matters. It will also mean better acceptance and more likely use of findings when the background is understood.

The entire group can contribute ideas to be worked into the evaluation plan and to be incorporated into the statement of purposes; leaders and fellow extension workers can consolidate the contributions into the final plan. All can volunteer to help when needed with data collecting and processing and, later, to help spread knowledge of the findings. A passive contribution all can make is to supply, without hesitation or reservation, any relevant personal data that may be required for the evaluation.

3. **While Evaluation Is in Progress**—"How are we doing?" would be the main question asked during this phase. Folks may not care much about going out of their way to find out for themselves, but they like to have some idea of what progress is being made. Whoever has leadership responsibility for the evaluation must see that periodic progress reports are made so that interested persons can keep in touch, and not have the opportunity to forget all about the matter.

The difficulty of reporting satisfactorily increases directly with the size of the group concerned and inversely with the amount of involvement that can be maintained. The larger the group involved, the more urgent will be the demands on ingenuity for devising ways of keeping the members informed and involved. Key subordinates will have to assume ever larger responsibilities in this regard as the leaders will have to devote ever larger portions of their time to coordinating and expediting activities, to executing the study plan and to processing the data.

A reasonable number of fellow extension personnel and qualified laymen can be kept intimately involved with data collecting and processing and, in some instances, with analyzing and interpreting tasks. Others will have to be kept in touch by periodic announcements of progress so they will not have to be brought back "into the fold" when they can make their next big contribution by aiding with word-of-mouth dissemination of results of the evaluation.

4. **When the Evaluation Is Completed**—When we have assembled the information and there are some results to report, we have reached the fourth stage. Now we should be ready to deal with the question, "What did we find out?" If the evaluation is quite elaborate, it is a good idea to get out tentative or preliminary reports on useable information as soon as we can assemble it. This gives associates an opportunity to get to work on the implications and some likely applications before interest cools off, and to sustain interest until the final report is ready for distribution.

They can begin work on followup reports and can prepare reports, news releases and radio or television talks for immediate use. They can tailor reports for the different groups within their audiences and can incorporate the material into their teaching. When the final report is ready, they should have worked up such a thorough understanding of the elements of the study that they can devote their full attention to grasping the "big picture" as it relates to their work and responsibilities.

Insofar as the immediate group is concerned, the final report should be something of an anticlimax If members have been kept properly informed, they should be aware of most of the contents of the final draft. Their main use of the document should be for future reference or as an aid in explaining the evaluation to others. However, some will want to use it when documenting followup reports relating its findings to those of evaluations done elsewhere, or to suggest areas for future evaluations. Such uses are, of course, in addition to the basic one of presenting the facts in such a manner that decisions regarding change or direction of action can be made with confidence. In this capacity its period of heavy use will depend on reaction to contents.

5. **Afterward**—Stage five comes after we have completed the evaluation and have made our reports. The sort of question we must now ask ourselves is, "Who should be informed of our evaluation?" Our responsibility now becomes a matter of sharing with others not directly concerned in the original evaluation but who are in positions where the information we have assembled might be used to advantage.

We should always make it a practice to pass along our findings to the State extension training and research persons who will be happy to see that our information comes to the attention of others in the State, or in other States, who might be interested. If they have not already been involved in the evaluation, you may want to involve them at this stage by securing their assistance in helping you implement the action suggested by the evaluation.

This is no time to neglect the folks in the study area who may have been involved only incidentally up to this point. They should have an opportunity to see the final report to be assured that the evaluation was completed (that the evaluator wasn't just doing some busy work) and to get some idea of what comes next. It will help them gain background for the discussion, lectures, radio and television talks, demonstrations, and the like that will be coming along in the followup action, and will help them grasp more quickly the intent of the plan as a whole.

Summary

We opened this chapter with a resume of the three conventional ways of gaining answers to questions (improve extension educational practices): Subjective impressions, consultation with professional researchers, and evaluation. We suggested disadvantages of the first two which are resolved in the last way.

Next, we discussed the difference in orientation or motivation between the evaluator and the professional researcher, indicating that the former needs answers to specific questions and that the latter must strive for generalizing answers. We also suggested that the devices of the evaluator, at least when working on his own problems, help speed up if not eliminate the awareness phase, and bring the action phase into focus much quicker.

We closed with a discussion of ways to maximize involvement in the evaluation as an aid to successful eventual accomplishment of indicated action. We suggested things that can be done at each of five stages (while thinking about the need for evaluation, while preparing the evaluation, while performing the evaluation, when the evaluation is completed, and while the indicated action is being carried out) to help smooth out the process.

References

1. Corey, S. M. **Action Research to Improve School Practices.** New York, Columbia University Press, 1953.

2. Report of Work Group B. **Organizing for . . . Using . . . Research in Extension. Summary of an Extension Research Workshop,** May 6-10, 1957. Washington, D. C., Federal Extension Service, U. S. Department of Agriculture.

Chapter XIX

UNDERSTANDING AND USING EVALUATION REPORTS

Joseph L. Matthews

Our objectives in this chapter are to encourage extension workers to:

1. Appreciate evaluation and extension reports as sources of useful knowledge.
2. Acknowledge that an organized approach is necessary for efficient use of evaluation and extension research reports.
3. Learn how to evaluate a report and how to apply the findings reported in it.

Reports Must Communicate

At times we may lose sight of the real purpose of evaluation and research reports. Basically, a report is written by someone to communicate what he, usually the evaluator or research person, thinks is important for the reader to know about the subject on which he is reporting. Although there are certain things that the writer can do to make his communication more palatable, there are other things lying beyond his control that can interfere with his efforts to communicate. We will comment on both sides of this problem, although the former has been dealt with in Chapter XVII.

Inhibitors to Communication

Before going into ways of making the best use of reports, we might look at some of the factors that influence their effectiveness as a method of communication. All too frequently, readers of evaluation or research reports tend to be more critical of the research methods, and the ways they are reported, than is justified. This attitude arises in part from a feeling that the obligation to communicate is solely that of the writer.

Obviously, this is an unreasonable attitude. No writer can anticipate all the possible situations under which the reader might be able to, or want to apply the results of a particular study. The reader should be willing to make some of the modifications or interpretations necessary to relate the study findings to his own situation. Unfortunately, this feeling is fairly widespread. It interferes with much potential use of findings from evaluation and extension research.

Communication in evaluation, more so than in some other situations, must be a two-way proposition. Both the sender and the receiver of the message being transmitted must be **active** participators in the communication process. The writer must be able to send clearly and the reader must be willing to receive the message and then adapt it to his own use.

Aside from certain inherent limitations of the written word, there are definite limitations on what the writer can do to make it easy for the reader to get the findings quickly and effortlessly. The pressure to simplify involves some danger of misleading the reader. It is a rare research person or evaluator who is personally familiar enough with all possible applications to apply the findings to all possible situations. You must have knowledge of the practical situation of the moment to make specific applications.

Likewise, a person who does not fully understand the study cannot safely make applications to practical operating situations which differ in any way from the test situation. If the evaluator and the program person are not the same individual, then the best substitute is for the evaluator and the program person to work very closely as a team while making interpretations and working out practical applications of findings.

Options and Obligations

When it comes to writing his report, the evaluator has three options. First, he can describe the situation in which the investigation was done in sufficient detail so that the reader can draw his own inferences about usefulness of findings for another situation. A second choice is to make qualified generalizations that still leave ultimate use decisions to the judgment and imagination of the reader.

The third choice is to describe specific situations, either actual or hypothetical, in which the findings would apply. Then the reader can decide on the degree to which he considers his own situation to be comparable to the sample situations, and can make his own decisions on how widely to apply the evaluation findings.

This gives the writer quite a range of methods of presentation. However, it gives the reader an equally wide range of expectations. When presentation and expectation patterns are not in agreement, the reader may be displeased with the report. This could be an argument for several presentations as we suggested in Chapter 17, but it could also be the occasion for a little more tolerance on the part of the reader.

At the same time, limitations of this variety do not excuse the writer from doing his best to write for the particular readers he has in mind while writing the report. This includes the requirements of writing as clearly and as simply as possible, using terms and examples that are within the experiences of the readers, using appropriate illustrative and explanatory materials, and so on. Finally, the evaluator should try to give the reader the benefit of his own unique interpretations, which he alone can make because of his thorough involvement in the study.

Personalized Presentations

The written report alone must be regarded as a less than satisfactory method of disseminating findings of evaluations or research studies. Ideally, as we suggested in the previous chapter, the possible users of the results of an evaluation should have the added benefit of face-to-face discussions with the evaluator. Full benefit from any scientific inquiry cannot be achieved without the stimulating influence of personalized communication processes. Whenever possible, therefore, you should plan to combine both written and personalized presentations when you desire to achieve a maximum flow of communication and understanding.

Brevity

Another unrealistic expectation held by some is that brief summaries of findings are adequate substitutes for full reports. The value of this technique for reporting diminishes as the complexity of the behavior being studied and the number of findings to be reported increase. Some things are so involved that brief summaries or broad generalizations are likely to be more misleading than informative.

These paragraphs are intended to demonstrate that written reports have definite limitations and to suggest the nature of their limitations. The purpose of the remainder of this chapter is to offer some suggestions to help you make good use of written reports in spite of any limitations they may have. The suggestions that follow can aid you in making systematically effective use of reports on evaluation research.

Three Operations

Using reports involves three main operations, each of which has several important steps. The three operations are, (1) evaluating the quality of the investigation that is reported, (2) extracting the significant findings, and (3) applying the findings to your own situation.

Your attention may be drawn to a report in any of several ways. The report may come to your desk through the mail; you may come across it while seeking information on a related topic; someone may call your attention to it to ask your opinion about something mentioned in it.

The first questions for you to answer to your own satisfaction is, "Would this study or research project be likely to produce results that could have bearing on my work?" Another question you must answer is, "Were the design and the methods employed in the study sound?" The answer to this question helps answer the next one: "How much confidence can I have in any findings that are reported?"

These questions are aimed at something quite different from the report itself or the interpretations made by the writer. Their purpose is to find out if the study is conducted in a way that is scientifically sound so that only facts are brought out and so little or no personal bias influences the outcome. No amount of statistical precision or careful interpretation can correct inaccuracies introduced through poor design or faulty methods of obtaining information.

Appraising the Quality of a Report

To do an adequate job of appraising the quality of a report you need satisfactory answers to questions in five main areas. The kind of information you need can be summarized under these headings.

1. **Purpose**—"What did the investigator try to do?" The important thing here is whether or not the questions he asked or the hypotheses he tested are ones in which you are interested. If you are not interested in them, or see no immediate likelihood of being interested in them, there is little point in your going further into the report.

2. **Method**—"What kind of a study was done?" and "How was it done?" are the questions under this heading. The decision here is whether or not the general design and the techniques employed were such that valid and reliable information of an objective nature can be anticipated. The term **valid** is used to mean really measuring what was supposed to be measured; reliable is taken to mean that the same kind of information would be obtained if the methods were repeated on other samples

of the population. As explained earlier, **objective** means that information or findings are not influenced by the personal point of view of anyone.

3. **Population** — "Exactly who were the people from whom or about whom information was obtained?" You will need to know this to determine how comparable the test situation is to the one to which you desire to apply the findings. While deciding the degree to which the people involved in this particular study are like any group of people you are concerned with, you will be trying to find in the description at least some of the same characteristics that are found in your own group.

In most cases you will find some differences or that certain characteristics that you consider important are not mentioned in the report. In that event, you will have to make personal judgments or estimates regarding the characteristics not mentioned to assign their probable bearing on the findings of the study before you can make use of the findings.

4. **Sample**—"What was the relationship of the sample to the whole population?" and "How was the sample obtained?" are the related questions asked at this point. The sample must remain a matter of concern, even after you are convinced that the population is quite similar to your own. A poor sample can so distort the reported characteristics of the population that the study might as well be based on an entirely different population. All too frequently, information about sampling methodology is not included in the published report, with the result that many unjustified assumptions of comparability are made.

5. **Timeliness**—"How recently was the information obtained?" and "Does it apply today?" Favorable answers to these questions cannot be made without considering many related factors. A recent date of publication or recent inclusive dates of field work do not guarantee data which will serve your purpose. Rapid changes in conditions of rural living, and in technological knowledge related to all aspects of contemporary life, force us to weigh the appropriateness of the findings in terms of detailed aspects of the individual situation rather than in broad terms of temporal or geographic similarity.

Date of publication is only a partial check on timeliness. The study may have been completed quite some time in the past and delayed months or years because of prior publishing commitments, budgetary problems of the sponsoring agency, or editorial delays of one kind or another. Not infrequently, studies are based on research completed many years previously or on data sources, such as the census, which may have been assembled ten or more years before the time they are used in the study.

Summary

"Look Before You Leap," might well be an alternate title for this chapter. The major emphasis of the last half is on ways to avoid hasty decisions regarding comparability and assumed applicability to your own situation. We have suggested questions in five areas to which you should be able to answer favorably before you should be willing to accept the findings of a study as appropriate to your needs.

Our objective in the earlier portions of this chapter is to make the point that the primary responsibility for communicating may be assigned to the writer, but that the reader has an obligation to try to understand what the writer is attempting to communicate. Even when a study, or the reported information about it, may have limitations, there may be many ways that the findings can serve as a contribution to knowledge if proper interpretations and allowances are made by the reader.

Chapter XX

IN CONCLUSION

Ward F. Porter

Evaluation is a process of systematic appraisal by which we determine the worth, value, or meaning of **something.** This **something** in Extension may be a **program** or part of one, a **method** or approach used in carrying on extension work, or a **situation** such as a community, a county, or even a larger area. Evaluation is a process that facilitates effective decision-making.

Why We Need Evaluation

The ultimate goal of the evaluation process in Extension is the increased effectiveness of our programs. Evaluation helps in this respect by telling us how much progress our clienteles have made toward their educational objectives. The process provides us not only with the results of our efforts, expressed in terms of changes in behavior, but also the benchmarks by which this progress can be gauged. In addition, evaluation contributes to our educational efforts by giving us necessary information about situations for planning sound programs.

In the increasingly complex world in which Extension functions, the evaluative process assumes growing importance and significance. Extension's responsibilities now extend far beyond the farm horizons of yesterday. Our clienteles are now found in cities and towns, as well as in the open country. Our audience is better educated, better informed, and more sophisticated.

Because of the technological revolution in communications and transportation, people can be more easily contacted through various channels, by mass media, for example. At the same time people can be influenced by agencies and organizations other than Extension through these same channels.

Because of these, as well as other changes in our society, there is grave need for more systematic evaluation in Extension. No longer is it enough to know that farmers apply fertilizer. Now we should know whether they are using the right kinds and amounts, their attitudes toward this practice, and various other more or less qualitative items of information. Likewise, since there are many sources of information currently available to farmers, we must be more zealous than ever in determining Extension's contribution to whatever change has taken place in our clientele.

The need to emphasize evaluation in Extension is evident. The techniques by which this process is carried out are less widely understood. Like any other skill, evaluation involves a fund of knowledge and an appreciation of certain basic procedures and "know-how." We hope this manual has made it clear that any extension worker can develop this knowledge and, with some experience, the necessary skill to carry on effective evaluation. This has been our premise in preparing this manual.

Basic Elements of Program Evaluation

Some of the basic techniques and procedures by which surveys and formal evaluations are conducted have been discussed in the preceding chapters. Expressed in much more summary fashion, we can visualize program evaluation as consisting of four procedural elements: (1) A statement of specific educational objectives; (2) an inventory of whatever was done in the name of the program to achieve the objectives; (3) an appraisal of the situation before the start of the program—that is, benchmark determination; and (4) subsequent measurement of the situation to determine progress and accomplishment. These can be thought of as more or less consecutive phases or steps in evaluating any given program.

Pin-Pointing Objectives—The first step in any program evaluation calls for an appraisal of specific program objectives. In this instance, we must answer the question, "What were we trying to accomplish?" The "What" must be **specific** enough to be measurable; and it must specify the changes in the situation (i.e., in the people) that the program hoped to accomplish. It must also indicate the particular audience involved.

Program Activities and Methods—In this, the second phase, we must account for all actions taken in the name of the program. This is not always a simple task, particularly where many agencies and organizations are involved. Because of our failure to do this adequately, we in Extension are sometimes in the uncomfortable position of being credit-ed with, or appearing to claim credit for, accomplishments for which we should not legitimately expect credit.

This difficulty can frequently be overcome, providing there is close cooperation between both program personnel and the evaluators. This is one instance where the right hand **must** know what the left hand is doing.

Establishing Benchmarks — One of the most essential steps in any program evaluation is an accurate appraisal of the situation before the program starts. Unfortunately, this appraisal is frequently neglected until sometime after the program is under way. We recommend appropriate benchmark surveys at the beginning of all programs. In any event, there has to be some evidence of what the pre-program situation was like before we can decide how much progress has been made.

Determining Program Progress and Accomplishments—The fourth basic step in evaluation is to determine progress or end results. Findings or conclusions of this nature must obviously be interpreted on the basis of program objectives—i.e., the anticipated educational changes. They must also be expressed in terms of the pre-determined benchmark. Further, since we are evaluating extension work, we must consider the activities and methods used in pursuing program objectives. Only in this way can we correctly and safely appraise the success of our efforts.

The Challenge

Evaluation is a tool; it is not an end in itself. It is the means by which extension programs and methods can be made more effective. If an evaluation is to be effective, the extension workers involved must incorporate the findings into their on-going program.

All extension workers evaluate. When they do so systematically and objectively, they derive increased personal satisfaction from the sure knowledge that their work has achieved some measure of success.

APPENDIX

Appendix A

ILLUSTRATIVE EVALUATION OUTLINE EVALUATING A PASTURE IMPROVEMENT PROGRAM

Fred P. Frutchey

The following is an illustrative outline for evaluating the results of teaching in carrying out part of an extension program. The main purpose of this evaluation is to find out what degree the teaching on improving pastures has been effective. To what extent were the objectives of the teaching efforts reached?

The decision to be made on the basis of the evaluation was, "Should the emphasis on pasture improvement be continued or are pasture improvement practices well enough established that time and effort can be shifted to other pressing problems?" This outline consists of ten steps:

1. Background Situation—The need in the county out of which the pasture improvement program grew.
2. Program planning—The procedure by which the problem got into the program and plan of work.
3. Objectives of the Program—The results expected from the program.
4. Extension Teaching Methods—This shows what was taught and how.
5. Questions to Be Answered by the Evaluation—These are the specific questions to be answered.
6. Sampling the Population—This includes the people to whom the pasture improvement program applies and the sample.
   a. Defining the population.
   b. Selecting the sample.
7. How to Get the Information—This step is concerned with the procedure for obtaining the information needed.
8. Preparation of Questionnaire—This is the form on which the replies of those who answer the questions are recorded.
9. Tabulation, Analysis, and Interpretation of the Information—In this step the information on the questionnaires is analyzed, organized and summarized to determine what it means for the program.
10. Use of the Findings—The application of the findings to program adjustment.

Steps Illustrated

1. **Background Situation**—The need in the county out of which the pasture improvement program grew. In Green County, many farmers have been changing from row crops to livestock but their pastures are inadequate. Among other things the pastures need lime, fertilizer and reseeding. The county agent had worked on this with some individual farmers but no concerted effort had been made over the county.

2. **Program Planning**—The procedure by which the problem got into the program and plan of work. This situation was presented at the annual meeting of the county extension program planning committee pointing out that the above practices were among the important ones recommended by the experiment station to build better pastures. After some discussion the committee selected these practices for emphasis in the program. The county extension staff was advised to spend considerable time on this phase of the program.

3. **Objectives of the Program** — The results expected from the program. From their knowledge of the situation and on the basis of the recommendations of the committee, the county extension staff formulated the following objectives of the program. These were the changes they hoped to bring about in the farm families in helping them to help themselves.

The objectives should be clearly and definitely stated. Vague objectives must be analyzed until they are observable evidences and made tangible. They may be what people know, do, or their attitudes.

The county extension staff decided upon the following objectives of their teaching program.

During the next three years to get at least 400 of the 500 livestock farmers to improve their pastures, thereby making it possible to support more livestock. Four specific objectives were set down. They were the three practices stated earlier and an additional practice—testing the pasture soil; to get the livestock farmers:

a. to lime their pastures
b. to fertilizer their pastures
c. to reseed their pastures with recommended seedings
d. to test pasture soil

These, then, are not only the objectives of the teaching program but are also the results to look for in the evaluation.

4. **Extension Teaching Methods**—This shows what was taught and how. These two columns show what was done to reach the objectives of the program. In the second column are the learning experiences provided to influence farm families to change.

WHAT WAS TAUGHT

a. Liming and fertilizing would produce a more luxurious growth of grass.
b. Reseeding will produce a more complete stand of grass and utilize the acreage more fully.
c. These practices are recommended as the result of research at the experiment station.
d. Some farmers in the county have tried them out with success, richer pastures supporting more livestock.
e. How they can get and apply lime and fertilizer, the use of soil tests, and how to get their soil tested.

HOW IT WAS TAUGHT

a. Circular letter sent to all livestock farmers in the county calling their attention to this situation and inviting them to come to a meeting in their part of the county.
b. Meetings in various parts of the county, using visual aids, and testimonies of farmers who had adopted the practices.
c. Demonstrations on farms in various parts of the county.
d. Pamphlet on improving pastures.
e. Timely news articles and circular letters.
f. Three major radio talks and short references from time to time on the radio.
g. Exhibit at the county fair.
h. Posters at lime, fertilizer and seed dealers.
i. The home demonstration agent mentioned and talked about these practices at home demonstration club meetings. In 4-H Club work the practices were discussed, particularly in the livestock clubs.

5. **Questions to Be Answered by the Evaluation**—These are the specific questions to be answered. The questions to be answered were, of course, the degree to which the objectives were reached (the results) plus any other pertinent results that may be expected. These questions were formulated as the ones on which information would be collected.

a. How many livestock farmers limed their pastures?
b. How many livestock farmers fertilized their pastures?
c. How many livestock farmers reseeded their pastures?
d. How many livestock farmers had their soils tested?
e. How many farmers increased their number of livestock?
f. What kind of livestock were increased and how many of each kind?
g. What was the estimated increase in value of livestock?
h. How did this compare with the cost of the adoption of the practices?

These and perhaps other questions would be answered by the evaluation. A questionnaire would be prepared to get answers to the first six questions. The analysis of the replies would consider the farms on which the practices were adopted for the first time during the three year period and the livestock increase per acre.

6. **Sampling the Population**—This includes the people to whom the pasture improvement program applies and the sample for the evaluation.

Defining the Population—The population was, of course, the 500 livestock farm families in the county. These are the people being taught. They are the ones at whom the program was aimed and who are in a position to do something about the practices.

Selecting the Sample—You want to get the answers to your questions as economically as possible (least time and effort) and with as much accuracy as you need. A sample of 100, **chosen at random,** from the 500 livestock farmers will give fairly accurate results and at the same time reduce the amount of time and effort spent on the evaluation.

There are several ways of selecting a sample. In this case every fifth name on the office's alphabetical list of livestock farmers in the county was selected for the sample. Selecting at random does **not** mean selecting haphazardly. It means a systematic method without subjective bias.

7. **How to Get the Information**—This step is concerned with the procedure for obtaining the information needed in answering the questions.

The information about the practices could be obtained by sending a mailed questionnaire and covering letter to the sample of 100 livestock families. This looks like the easiest and quickest way but often is not the case. Many families don't return the questionnaires without one or more follow-up letters which extends the time for getting the information and increases the follow-up work. You must have all or practically all of the 100 questionnaires returned; otherwise your sample may become biased and the error increased.

Another method is to interview each of the 100 livestock families in the sample. This method sounds at first glance to be too cumbersome and time-consuming. On further consideration it can be made quite practical. The three agents in the county could divide up the 100 families; each agent having 33 or 34 families to interview, taking about 3 days or so. The interview not only serves to collect the information in a very reliable way; it also serves as a helpful farm and home visit and as a follow-up of a program, thereby showing the interest of the agents which would be appreciated by the families.

It could also be done by members of the county program planning committee after they have been properly trained to do so. This has certain decided advantages for the type of evaluation used in this example.

8. **Preparation of Questionnaire**—This is the form on which the replies of those who answer the questions are recorded. It may be a questionnaire to mail or for use by interviewers who visit the farm for the information. In any case it is a form on which the information is recorded. The example which follows was partially designed as a mail questionnaire.

Questionnaire (Example)

PASTURE PRACTICES IN GREEN COUNTY

(Date)________________

Dear Sir:

This is a mailed questionnaire sent to 100 livestock farmers selected as a cross section sample in Green County to survey the pasture practices used.

Your reply will be greatly appreciated since the reliability of the survey will depend upon having every questionnaire returned.

We believe the questions can be easily answered in about 10 minutes.

We will appreciate your cooperation very much.

Sincerely yours,

(signed by the county agent)

1. Do you lime your pastures? (Check)
a._____Yes; b._____No.
2. If Yes, when did you start to do so?
a._____During the past 3 years.
b._____Before 3 years ago.
3. Did you do more liming of your pastures during the past 3 years than you did previously?
a._____Yes; b._____No.
4. Do you fertilize your pastures?
a._____Yes; b._____No.
5. If Yes, when did you start to do so?
a._____During the past 3 years.
b._____Before 3 years ago.

6. Did you fertilize your **pastures** more during the past 3 years than you did previously?
a.______Yes; b.______No.

7. Do you reseed your pastures? a.______Yes; b.______No.

8. If Yes, when did you start to do so?
a.______During the past 3 years.
b.______Before 3 years ago.

9. Have you reseeded your **pastures** more during the past 3 years than you did previously?
a.______Yes; b.______No.

10. Have you had your **pasture soils** tested during the past 3 years? a.______Yes; b.______No.

11. How many of each kind of livestock do you have on your farm?

	Now (1)	**Three years ago (2)**
a. Dairy cattle		
b. Beef cattle		
c. Sheep		
d. Hogs		

12. How many acres do you have in pasture?
a.____________Now; b.____________Three years ago.

13. Have you heard about the pasture improvement program in the county? a.______Yes; b.______No.

9. **Tabulation, Analysis and Interpretation of the Information**—In this step the information on the questionnaires is analyzed and summarized to determine what it means for the program.

When the questionnaires are completed, one or more of the agents can edit them, prepare a "typewriter tabulation" sheet and show a secretary how to tabulate the replies. These will be analyzed and summaried by the agents and a report prepared. The following is an example of "typewriter tabulation" sheet.

10. **Use of the Findings**— The application of the findings to program adjustment. The report will be presented to the county program planning committee for its study and discussion and to determine the extent to which "improving pastures" should continue to receive emphasis in next year's county extension program. The results of the evaluation will give the agents some clues as to the effectiveness of the teaching methods used.

The information will serve as a topic for a circular letter, radio program or news letter which has certain public relations values. The agents will have the satisfaction of knowing very definitely the degree to which the objectives of that phase of the program were accomplished and of their contributions in the county. The following are examples of tables and charts.

TABLE FOR REPORT (Example)

Percent of Livestock Farmers Using Improved Pasture Practices

Pasture Practices	**Percent of Livestock Farmers**
1. Liming	
a. Lime pasture now	80
b. Started liming during past 3 years	60
c. Did more liming during past 3 years	70
2. Fertilizing	
a. Fertilize pastures now	75
b. Started fertilizing during past 3 years	70
c. Did more fertilizing during past 3 years	75

Total Questionnaires Received—100

County__________________

TABULATION SHEET (Example)

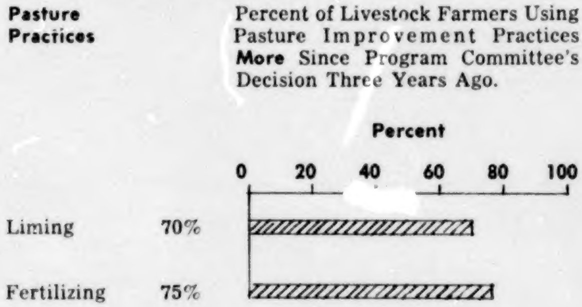
Questionnaire number and/or name	1. Lime pasture			2. Started liming			3. More liming during past 3 years			4. Fertilize pasture			5. Started fertilizing			
	a. Yes	b. No	NR	a. Past 3 yrs.	b. Before	NR	a. Yes	b. No	NR	a. Yes	b. No	NR	a. Within past 3 years	b. More than 3 yrs. ago	Don't know	NR
1	X			X			X			X					X	
2		X								X					X	
3	X			X			X			X			X			
4		X									X					
5	X			X			X			X			X			
This illustrates the tabulating procedure. Count all the X's in column 1a and write that number at the bottom of column 1a.																
Total	80	15	5	70	20	10	70	20	10	75	25	0	60	15	15	10

---

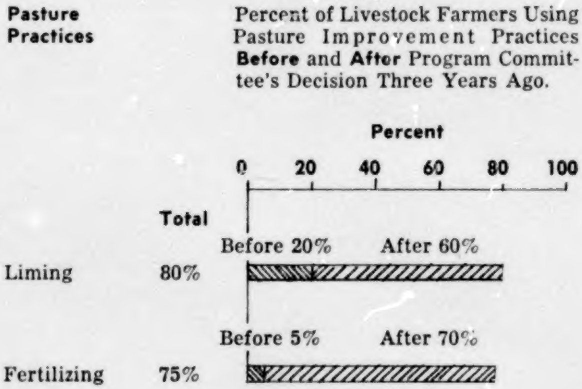
1. NR means "no response"

CHARTS FOR MEETING OF COUNTY EXTENSION PROGRAM PLANNING COMMITTEE (Example)

Percent of Livestock Farmers Using Pasture Improvement Practices **More** Since Program Committee's Decision Three Years Ago.



Percent of Livestock Farmers Using Pasture Improvement Practices **Before** and **After** Program Committee's Decision Three Years Ago.



Appendix B

EXAMPLE OF AN INFORMAL EVALUATION

Mrs. Laurel K. Sabrosky

This example of an informal evaluation illustrates the importance of planning ahead of time and involving learners to assure fullest interest and maximum benefits.

Evaluating a Series of In-Service Training Meetings

The audience at these meetings was about 30 or 40 extension workers; the speakers and discussion leaders were authorities in the various subject-matter areas; the planners were extension workers with training responsibilities. This is the procedure that was followed:

1. Each speaker and discussion leader was asked to provide the planners, before the meetings started, with a statement of the educational objectives he had for his part of the program, in terms of what he expected the learners (extension workers) to get out of it. The speakers and discussion leaders were completely cooperative in providing this information.

2. Near the end of the series of meetings, the audience was formed into working groups, and each group discussed and decided on the most important things it had learned from each talk or discussion.

3. The evaluation discussion leader placed on a blackboard the things learned by the audience from each talk or discussion. He placed these on one half of the blackboard. Then he wrote on the other half of the board the expectations of the speaker or discussion leader.

4. The total audience then reviewed the similarities and differences between what the speaker or discussion leader expected and what they thought they had learned. All speakers and discussion leaders who were not involved in other professional duties at the time were present for this evaluation session.

5. Members of the audience discussed what they were expected to learn, why they had learned what they were expected to learn from some speakers and discussion leaders, and why they did not learn the expected things from others. The evaluation leader brought out that they were following the approved method of evaluating an educational effort: (1) Doing so in terms of the objectives of the "teacher," (2) involving the "learners," (3) analyzing the results to see reasons for success and failure, and (4) looking ahead to occasions for applying the results to similar situations.

This evaluation procedure not only illustrates a means of evaluating properly, but also provides a method of reemphasizing the points the audience has learned, and of bringing to them points that they have missed.

Appendix C

PROFESSIONAL HELP FOR THE EVALUATOR

Joseph L. Matthews

The potential area from which to select an evaluation project in extension teaching is as broad as Extension itself. Any project you might select could deal with any one phase, or a complex of phases, which could range from one extreme of extension activity to another. At the same time, no extension worker can hope to be an authority in all areas of the work. Therefore, any of us who wish to pause and assemble information that will help us improve our work need have no feeling of embarrassment over not having all necessary knowledge and skills to follow through with appropriate evaluation.

For example, many of us are trained in special areas of agriculture or home economics. Once we are on the job, however, we find that many of our administrative and coordinating tasks require considerable understanding of the social sciences. We need to know techniques for identifying differences between persons who should or could be reached and those who actually are reached by our teaching methods. We need to know why certain groups among our clientele do not respond so well to our programs as we believe they should. We need to know the differences between urban and suburban families in our county, and how these differences may influence their demands on our services.

We need to know how best to approach certain families or individuals—how to work through their needs, values, motivations, ambitions, knowledge and understanding. We need to know which methods will work most effectively with which groups—how to approach them, how to make them interested, how to hold their attention, how to encourage them to follow through to completion, and how to carry on by themselves.

Although these are everyday concerns with most of us, it is easy to see that gaining the necessary information in the proper form to make the kind of evaluation that would help us in our work would require more than routine understanding of evaluation methods and techniques. In spite of the certain value such knowledge would have for us, the size of the project makes us pause to consider the question, "Should I undertake evaluation activities that are beyond my present competence?"

Such a question implies that we feel we have to "go it on our own" in evaluation. Actually, there are many professional resources which are at our disposal. The kind of question we should be asking ourselves is, "Are the needed resources and technical competences available to help with the evaluation I want to make?"

When we state the question in this manner, it is much easier to gain the confidence to start for, in most instances, we will find that all the professional help we will require is at our disposal. All we have to do is look around and take counsel where needed from persons who are competent in the areas that are difficult for us or unfamiliar to us. With their help, we can move on to appraising what resources we will need and finding out whether or not they are available.

The final decision to undertake an evaluation, and how extensive the investigation shall be, will depend upon the availability of technical help and other special resources that must be marshalled for the job. Technical help is essentially a personnel problem—a matter of contacting the right people and enlisting their knowhow. The other resources consist of time allocations by persons involved in the evaluation, and any special materials or facilities that may have to be assembled.

As in all investigative activity, the evaluation project finally outlined and executed will be a compromise between what we would like to do and what we were able to do. We find out sooner or later that others are not so convinced as we are as to the relative importance of the evaluation we contemplate and that they have other ideas for using limited available common resources.

Before we complete the evaluation task we can expect to make numerous compromises for reasons of economic or practical expediency. However, compromises provide practical educational experiences in resourcefulness and perseverence and should not be taken as reason for dropping a research or evaluation project unless they invalidate the basic objectives of our efforts.

If you have never been involved in a systematic evaluation activity, you may be at a loss as to where to turn for the help you need. This confusion should not last long, however, as we in Extension are fortunate in that we have ready access to a wide variety of professional guidance and help.

All of us have access to the technical and professional resources of our own land-grant institution, and some of these resources are exceptionally helpful. Of course, there will be some variation in the particular assortment of competences represented in a particular State extension staff, resident teaching staff, or experiment station staff, but you can count on finding someone who can help you get a start, or who can tell you where to look if he does not feel qualified to advise you.

Availability of such persons may be limited by other demands placed on their time, but most of them can find time to get you started with your evaluation. Another factor which may interfere with their effectiveness insofar as your work is concerned is the distance from the institution to your study area. If you can run in or your consultant can come out in a few minutes, you can get together about as often as you care to. If the trip is one which will require hours or days, both of you will have to plan your meetings carefully, and work up agendas for when you can get together. At other times you will have to resort to correspondence or to working with substitutes he may be able to recommend.

Many times he can suggest specialists and facilities of his institution which are located off-campus but which may be more convenient to your study area. There are many facilities of this nature which you can feel free to draw upon. No doubt you will be amazed when you learn how many of them there really are. For the balance of this appendix section, however, we will confine our remarks to facilities available from the land-grant institutions and the Federal Extension Service.

State Extension Staff

You are no doubt aware of the many specialists on the Cooperative Extension Service staff of your State but you may not be aware that over 20 States have a specialist in extension research on the staff. Your State may be one of them. This staff member is the proper initial contact for aid in any kind of evaluation project. In most instances he has had special training in the techniques of evaluation and is quite likely to be highly competent in social science research.

He can assist you personally in planning and following through on an evaluation activity. In addition to the help he can give you personally, he is in touch with other individuals in resident research and teaching who have the skills you may find it necessary to draw upon.

Most of the 30-odd State personnel training specialists have some competence in evaluation and research in general. In the absence of a State extension research person, the leader of the training program is a good person to contact.

Most State staffs have one or more persons who have participated in formal evaluation work, or who have taken the evaluation course at a regional extension summer school. If there is no extension research or training person on the staff, a person with this kind of background can be consulted.

Resident Teaching Staff

All resident teaching staffs at the land-grant institutions have individuals with backgrounds of research in education or the social sciences. Usually there is someone on the staff who has had research experience directly related to almost any conceivable evaluation problem we might have. All you have to do is find the proper person and make suitable arrangements to get the help you need from him.

Good persons to seek out when you need help on evaluation are the ones teaching research courses or those engaged in research work in the various subject-matter departments. The more experienced staff members in departments of mathematics or statistics, psychology, education, economics, sociology and anthropology can make helpful suggestions regarding evaluation methodology. Although their suggestions would have to be modified to fit your particular needs, they should be helpful in extending your understanding of general approaches and basic investigative techniques.

Experiment Station Staff

Every experiment station has on the staff persons whose work is planning and following through on a wide variety of research projects that involve essentially the same problems of design and execution that we encounter in our evaluation projects. This is particularly true of the people engaged in sociological research, sampling methodology, questionnaire design, and data collecting and processing.

Most experiment stations have specialists who devote most of their time to working out coding plans for questionnaires and schedules for either hand or machine tabulation. They also serve as consultants on questionnaire and schedule designs. Usually there is someone who is experienced at preparing tabulation plans in sufficient detail so that the followup work can be done by clerks or machine operators. All of them are available to advise you in processing your own data for analysis and interpretation.

Most of the experiment stations have modern electrical machines for tabulation and other analysis operations to speed up data processing. In many States the facilities are adequate to handle almost any statistical job likely to be undertaken at the institution. During slow periods they are made available to outside researchers. At such times you may be able to arrange for use of the machines, together with the services of the operators, to process your data.

Adequacy and availability of facilities and qualifications of staff members and technicians will vary from one institution to another, but all of them will be able to help a great deal. Most of the time the resources are more than adequate for routine evaluation work. This is particularly true of the kind of evaluation we would do by ourselves or with the cooperation and assistance of a few colleagues.

Even overloaded resident teaching and experiment station personnel are willing to advise us to the best of their ability when we approach them through proper channels. When they are convinced we are sincere in our efforts at evaluation, they will be most tolerant and understanding of our comparative incompetence, and will go out of their way and make personal sacrifices to be helpful.

Division of Extension Research and Training

This Division of the Federal Extension Service is a service unit that has as one major responsibility, the national leadership in extension research and the staff training in evaluation. This combined function is accomplished by coordinating extension research and evaluation, by training and assisting State research persons with their work, by providing courses and training sessions in evaluation, and by consulting and participating directly in evaluation activities at all levels.

The size of the staff and the facilities available to it limit active participation in evaluation activities to a restricted number of major national, regional, State and county studies. Most of these are done as cooperative projects between the State or regional group and the Division. All of them must be of the pioneer variety to some degree, although a certain amount of replication occurs.

The major contribution of staff members is made through personal consultation in the States or through correspondence with individuals to advise them on evaluation activities and extension research projects. Assistance provided may consist of giving suggestions about sources of needed help, supplying materials, advising on specific phases of the work, or reviewing plans for all stages of the contemplated evaluation sequence.

Requests for assistance from Division personnel by workers in a State should be made through the State Director to the Director of the Division of Extension Research and Training or to a particular staff member of the Division. If there is some question as to the identity of the person who can give the desired help, the request should be made to the Director who will refer it to the proper staff member. Otherwise, the State Director can communicate directly with the proper staff member.

The general categories of help that can be given by the Division are listed as major topics in the accompanying outline. Under each major topic are listed the specific kinds of assistance that are offered in that area.

1. Defining the problem and determining necessary resources.
    a. Spelling out the problem and the general nature of the information that will be needed.

   b. Investigating resource needs and possible ways of meeting them.
   c. Analyzing the time requirements and preparing a preliminary calendar of work.
   d. Analyzing the nature and extent of the cooperation that will be needed from other extension workers.
2. Designing the evaluation plan.
   a. Spelling out what was taught that is to be evaluated.
   b. Defining what aspects of the teaching action are to be evaluated.
   c. Deciding upon objectives of the evaluation.
   d. Deciding what evidence is to be collected.
3. Preparing research forms.
   a. Deciding what questions to ask and how to phrase them.
   b. Developing the record format.
   c. Preparing instructions for handling the questions.
4. Planning tabulations and the analysis.
   a. Preparing the coding plan.
   b. Deciding on the tabulation method.
   c. Preparing tabulation forms.
   d. Deciding on statistical techniques.
5. Preparing the sampling plan.
   a. Defining the population.
   b. Determining the sampling fraction.
   c. Preparing the sampling instructions.
6. Conducting the field survey.
   a. Selecting interviewers or recorders.
   b. Training interviewers or recorders.
   c. Supervising field party.
   d. Editing schedules.
7. Preparing the report.
   a. Planning the report outline.
   b. Analyzing the data and drawing implications from them.
   c. Writing the report.
   d. Publishing the report.
8. Getting findings disseminated and used.
   a. Interpreting the findings.
   b. Drawing implications for action programs.
   c. Planning the disseminating activities.
   d. Setting up reporting situations.

In addition to the direct help that members of the Division staff can give, there are numerous duplicated materials on evaluation in Extension that are available for distribution. These materials consist of examples of questions or complete questionnaires, various kinds of evaluation check sheets, tabular layouts and recording devices, and brief statements on methodology for the different phases of evaluation. They also have samples of different kinds of completed evaluation reports.

Staff members are prepared to answer most routine requests for suggestions about how to handle specific evaluation problems. In many cases they can provide appropriate example materials to help make their suggestions clear. When the problem is a new one to them, they will do all in their power to work out a practical solution.

Appendix D

EVALUATION REFERENCES IN EXTENSION SERVICE REVIEW

From time to time the Extension Service Review contains articles on the need for evaluation, planning evaluation, methods of evaluation, results of evaluation, and uses of the results. Some of these references are listed below.

January, 1955

1. Gallup, Gladys and Matthews, J. L.—Appraising Farm and Home Unit Approach. p. 19

April, 1955

2. Cowlitz County, Wash.—Who Are the Members of Home Demonstration Clubs? p. 84

May, 1955

3. Sabrosky, Laurel K.—Some Ways to Hold Your Young People. p. 101
4. Clayton, Leon O.—Factors for Successful 4-H Club Work. p. 105

September, 1955

5. Matthews, J. L.—What Is Evaluation? p. 171
6. Fifield, Marguerite and Whitman, George E.—After We Interviewed 346 Families We Revamped the County Program. p. 173
7. Knutson, Andie L.—How Do You Measure Progress? p. 174
8. Procter, Harriet—Our Time—Use Study Was an Eye Opener. p. 176
9. Harris, Margaret E.—We Take a Good Look at the Women We Serve. p. 177
10. Fessenden, Jewell—Tools for Testing. p. 178
11. Raudabaugh, J. Neil—Program, Plan of Work, Program Evaluation. p. 180
12. Ueland, Gale—Half of Louisville Listened. p. 183
13. Clark, Robert C.—The Home—Focus of Our Research. p. 184
14. Frazier, Leslie—Our Families Wanted Facts. p. 187
15. Schruben, Luke M.—Compounding State Research. p. 189

February, 1956

16. Pierson, Arlie A.—One Way to Measure the Value of Mass Media. p. 45

March, 1956

17. Blanchard, Evelyn and Shipley, Fern—Four State Staff Pool Ideas for a Revised 4-H Food Program. p. 55

May, 1956

18. Copp, James H.—Why Some Stay In and Others Drop Out of 4-H Clubs. p. 97
19. Kern, K. Robert—Iowa Folks Found the Answer to Facts—So What? p. 100
20. Marchant, John D. and Rowe, M. Pauline—Our Time-Use Study. p. 103

June, 1956

21. FES Task Force—Earmarks of Farm and Home Development. p. 109

July, 1956

22. Dewar, Carmen L.—Iowa Women Learn to Serve Nutritious Meals. p. 132
23. Crosson, Mary Alice—Indiana Specialists Test Their Teaching Skills. p. 134

August, 1956

24. Sevoian, Lucile Hieser—Complete—Not Complete; A Satisfying Project Experience in 4-H. p. 149

September, 1956

25. Ross, John E.—More Than Brick and Mortar. p. 157
26. Longsdorf, Lisle L.—Facts and Dreams. p. 163
27. Several Counties—Let's Tackle One Problem at a Time. p. 166
28. Wyoming Extension Specialists—Analysis of Needs —A First Step in Program Projection. p. 170
29. Swanson, Harold B.—Early Results of Program Projection. p. 172

November, 1956

30. Pope, Edward V.—Do You Know Yourself? p. 204
31. Beers, Howard W. and Rheinwald, Hans—Evaluation Takes Root in Europe. p. 206
32. Tichenor, Philip—A Hazard Hunt. p. 207

December, 1956

33. Dodge, W. A. and Bradfield, Alex—Vermonters Improve Milk Flavor to Increase Sales. p. 221
34. Stanislaus County, California—What Extension Help Do Women Want? p. 235

January, 1957

35. Erickson, Anna Jim—Listen to the People Speak. p. 11

March, 1957

36. Bergen County, New Jersey—Food—No. 1 in 180,000 Budgets. p. 63

April, 1957

37. Reeder, Robert—We Took Another Look. p. 77
38. Liveright, A. A.—What Makes Them Tick—My Observations of Two Successful Community Organizations. p. 79

May, 1957

39. Rutledge, Lloyd L.—Clues to 4-H Club Program Vitality. p. 105
40. Dildine, Glenn C.—Your Citizenship Improvement Study. p. 107
    Sabrosky, Laurel K.—Let's Be Challenged. p. 108

July, 1957

41. Holloway, Margaret E.—We Used the Facts We Learned From 500 Interviews. p. 139
42. Straus, Murray A.—What We Can Do for the Part-time Farmers. p. 153

October, 1957

43. Welch, Jr., C. Herman—Daily Records: Chore or Rich Harvest? p. 203
44. Barlow, Shirley—The Office Secretary—Envoy of Good Will. p. 206
45. Welch, Jr. C. Herman—Know Your County People. p. 210

December, 1957

46. Sabrosky, Laurel K.—Evaluation as Related to Public Relations. p. 252

January, 1958

47. Leagans, J. Paul—Give Your Fitness a Checkup. p. 6

March, 1958

48. Porter, Ward F.—Studying Our Human Resources. p. 57

April, 1958

49. Hill, Kate Adele—Three-Way Approach Pays Off. p. 79

May, 1958

50. Carter, Glenn—Wants and Needs of Older Youth. p. 108

September, 1958

51. Shipley, Fern—Probing to Determine Needs. p. 180

November, 1958

52. Sarbaugh, L. E.—Have You Tested Your Audience? p. 219

January, 1959

53. Frutchey, Fred P.—To Get . . . or Not to Get—Further Education. p. 3